Random matrix theory questions arising in Compressed Sensing and related areas

Bubacarr Bah
with Prof. Jared Tanner

School of Mathematics, University of Edinburgh
and
Maxwell Institute

Foundations of Computational Mathematics
July 2011
Budapest, Hungary
Compressed Sensing (CS) and Related Areas

Restricted Isometry Property (2) - RIP₂
- Restricted Isometry Constants (2) - RIC₂
- Improved RIC₂ Bounds for the Gaussian Ensemble
- Numerical Results
- Performance Guarantees for CS Algorithms
- Finite \((k, n, N)\) interpretations

Asymptotic Approximation of RIC₂ Bounds
- Approximation of Bounds
- Implications for CS

Restricted Isometry Property (1) - RIP₁
- Sparse random matrices and RIP₁
- Performance Guarantees for \(\ell_1\)

Conclusion
Compressed Sensing & Spare Approximation

- Signal $x \in \mathbb{R}^N$, $k$-sparse.
- Sensing matrix $A \in \mathbb{R}^{n \times N}$; measurements $y = Ax$, ($n \ll N$).
- Problem $(P^k_0) : \min_{x \in \mathbb{R}^N} \|x\|_0 \text{ s.t. } Ax = y$.
- Solution: $l_q$ minimizations & Greedy Algorithms (OMP, IHT, ...)

Rank Minimization & Matrix Completion

- Matrix $X \in \mathbb{R}^{m \times n}$, low rank $\text{rank}(X) \leq r$.
- Linear map $\mathcal{A}(X) : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^p$; measurements $y = \mathcal{A}(X) \in \mathbb{R}^p$.
- Problem $(P^r_0) : \min_X \text{rank}(X) \text{ s.t. } \mathcal{A}(X) = y$.
- Solution: $\| \cdot \|_*$ minimizations, & Greedy Algorithms (SVT, ...)

Bubacarr Bah with Prof. Jared Tanner
Random matrix theory questions arising in Compressed Sensing
Compressed Sensing (CS) and Related Areas

Restricted Isometry Property (2) - RIP_2
Asymptotic Approximation of RIC_2 Bounds
Restricted Isometry Property (1) - RIP_1
Conclusion

CS Applications

- Medical Imaging: MRI, fMRI, Radiology, ...
- Infrared spectroscopy & Seismic imaging
- Single pixel camera & Analog-to-digital converters
- DNA micro-arrays, radar, wireless communications, ...

Tools of Analysis

- Coherence  [Donoho & Huo; Elad & Bruckstein]
- Restricted isometry property  [Candès & Tao]
- Nullspace property  [Donoho & Huo]
- Stochastic geometry  [Donoho; Donoho & Tanner]
- Message passing  [Donoho, Maleki & Montanari]
RIP certainly a popular tool of analysis; FoCM’11?

With the introduction of RIP$_1$, we refer to the standard RIP as RIP$_2$ - the subscripts 1 & 2 refer to the norms used

**Definition**

RIC$_2$ of $A$ of order $k$ is the **smallest** number $R_k$, for all $k$-sparse $x$, such that

$$(1 - R_k) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + R_k) \|x\|_2^2$$

**Definition (Rank Minimization Equivalence)**

RIC$_2$ of $\mathcal{A}(X)$, the $r$-restricted isometry constant, is the **smallest** number $R_r$, for all matrices $X$ of rank at most $r$, such that

$$(1 - R_r) \|X\|_F^2 \leq \|\mathcal{A}(X)\|_2^2 \leq (1 + R_r) \|X\|_F^2$$
A having RIP$_2$ means that $A$ is a near isometry for $k$-sparse $x$

RIP$_2$ gives a sufficient guarantees for exact recovery

\[ \ell_1 \text{ minimization works if:} \]

- $R_{3k} + 3R_{4k} < 2, \quad [\text{Candès, Romberg & Tao, 2006}]$
- $R_{2k} < \sqrt{2} - 1, \quad [\text{E. Candès, 2008}]$
- $R_{2k} < 2/(3 + \sqrt{7}/4) \approx 0.4627, \quad [\text{S. Foucart, 2010}]$

Similarly for Greedy Algorithms:

- **IHT:** $R_{3k} < 1/\sqrt{3}, \quad [\text{S. Foucart, 2011}]$
- **CoSaMP:** $R_{4k} < \sqrt{2}/(5 + \sqrt{73}), \quad [\text{S. Foucart, 2011}]$
- **Subspace Pursuit (SP):** $R_{3k} \lesssim 0.06, \quad [\text{Dai & Milenkovic, 2009}]$
A more quantitative definition is the asymmetric RIP₂:

**Definition**

RIC₂ of A of order k is the smallest L & U, for all k-sparse x, s.t.

\[(1 - L(k, n, N; A)) \|x\|₂^2 \leq \|Ax\|₂^2 \leq (1 + U(k, n, N; A)) \|x\|₂^2.\]

**RIC₂ of A & eigenvalues of** \(A_K^*A_K\), for \(\Omega = \{1, 2, 3, \ldots, N\}\)

\[1 + U(k, n, N; A) := \max_{K \subset \Omega, |K| = k} \lambda^{\max} (A_K^*A_K)\]

\[1 - L(k, n, N; A) := \min_{K \subset \Omega, |K| = k} \lambda^{\min} (A_K^*A_K)\]

Thus L & U are smallest & largest deviation from unity of smallest & largest \(\lambda (A_K^*A_K)\) respectively.
• RIC$_2$ **combinatorial, intractable** for deterministic $A$, NP-hard
• Probabilistic bounds possible $\Rightarrow$ the use of random matrices
• Different approaches include:
  • Largest ensembles with bounded RIC$_2$, [Mandelson et. al.]
  • RIC$_2$ bounds for partial Fourier matrices, [H. Rauhut]
  • RIC$_2$ bounds for Gaussian matrices, [C. & Tao; B.C.T.]

**Goal:**
Calculate accurate RIC$_2$ bounds for Gaussian random matrices with entries drawn i.i.d. from $\mathcal{N}(0, 1/n)$

**Motivation:**
(1) Using the Gaussian to model **zero mean i.i.d** ensembles
(2) **Easier!** - Lot of literature available on the Gaussian ensemble
Linear growth or proportional-growth asymptotics (p.g.a)

Problem instances \((k, n, N)\) considered is where the following ratios converge to nonzero bounded limits:
\[
\frac{k}{n} = \rho_n \rightarrow \rho \quad \text{and} \quad \frac{n}{N} = \delta_n \rightarrow \delta \quad \text{for} \quad (\delta, \rho) \in (0, 1)^2 \quad \text{as} \quad (k, n, N) \rightarrow \infty.
\]

Theorem

Let \(A\) be a matrix of size \(n \times N\) whose entries are drawn i.i.d. from \(\mathcal{N}(0, 1/n)\). For any \(\epsilon > 0\), in the proportional-growth asymptotics
\[
P(L(k, n, N) < \mathcal{L}^{BT}(\delta, \rho) + \epsilon) \rightarrow 1 \quad \& \quad P(U(k, n, N) < \mathcal{U}^{BT}(\delta, \rho) + \epsilon) \rightarrow 1
\]
exponentially in \(n\).

- Prior bounds by Candès & Tao; and Blanchard et. al. (BCT)
**Derivation technique**: finding smallest $\lambda^{\text{max}} (\delta, \rho) > 0$ such that in the p.g.a, for $U_k = U(k, n, N; A)$,

$$
P(1 + U_k > \lambda^{\text{max}} (\delta, \rho)) = P\left( \max_{K \subset \Omega, |K|=k} \lambda^{\text{max}} (A_K^* A_K) > \lambda^{\text{max}} (\delta, \rho) \right) \to 0
$$

Candès & Tao used **union bounds and concentration of measure bounds** on the extreme eigenvalues of Wishart matrices - **valid for sub-gaussian matrices**

$$
P\left( \max_{K \subset \Omega, |K|=k} \lambda^{\text{max}} (A_K^* A_K) > \lambda^{\text{max}} (\delta, \rho) \right) \leq \left( \begin{array}{c} N \\ k \end{array} \right) P\left( \lambda^{\text{max}} (A_K^* A_K) > \lambda^{\text{max}} (\delta, \rho) \right)
$$

where $\lambda^{\text{max}} (\delta, \rho) := \left[ 1 + \sqrt{\rho} + (2\delta^{-1}H(\delta\rho))^{1/2} \right]^2$

Their **upper bound** is then $U^{CT}(\delta, \rho) := \lambda^{\text{max}} (\delta, \rho) - 1$
BCT achieved tighter bounds using union bounds and bounds of probability density functions of the extreme eigenvalues of Wishart matrices [A. Edelman, 1989]

\[
P(\max_{K \subset \Omega, |K|=k} \lambda^{\text{max}} (A_K^* A_K) > \lambda^{\text{max}} (\delta, \rho)) \leq \int_{\lambda^{\text{max}}(\delta, \rho)}^{\infty} \binom{N}{k} f_{\text{max}}(m, n; \lambda) d\lambda
\]

But \( f_{\text{max}}(m, n; \lambda) \leq p_{\text{max}}(n, \lambda; \rho) \exp(n \cdot \psi_{\text{max}}(\lambda, \rho)) \) where \( \psi_{\text{max}}(\lambda, \rho) := \frac{1}{2} \left[ (1 + \rho) \ln \lambda - \rho \ln \rho + 1 + \rho - \lambda \right] \)

Bounding \( \binom{N}{k} \) by the Stirling’s formula the exponent of the exponential term becomes \( \delta \psi_{\text{max}} (\lambda^{\text{max}} (\delta, \rho), \rho) + H(\delta \rho) \)

In the p.g.a only the exponential term matters and \( \lambda^{\text{max}} (\delta, \rho) \) becomes a solution to when the exponent is zero

Their upper bound is thus \( U^{BCT}(\delta, \rho) = \lambda^{\text{max}} (\delta, \rho) - 1 \)
 Improvement on BCT bounds achieved by grouping submatrices, i.e. for $A_K$ and $A_{K'}$ with $|K \cap K'| \gg 1$, hence decreasing the combinatorial term significantly

$$
P \left( \max_{K \subset \Omega, |K|=k} \lambda^{\text{max}} (A_K^* A_K) > \lambda^{\text{max}} (\delta, \rho) \right)$$

$$= P \left( \max_{i=1, \ldots, u} \max_{K \subset G_i, |K|=k} \lambda^{\text{max}} (A_K^* A_K) > \lambda^{\text{max}} (\delta, \rho) \right)$$

The RHS upper bounded using union bound over groups of $m \geq k$ distinct elements and controlling dependencies in $\lambda^{\text{max}} (A_K^* A_K)$ for $K \subset G_i$ by replacing the maximization over $K \subset G_i$ by $\lambda^{\text{max}} (A_M^* A_M)$, $M := \bigcup_{K \subset G_i, |K|=k} K$, $|M| = m \geq k$.

$$\text{RHS} \leq u \cdot P \left( \lambda^{\text{max}} (A_M^* A_M) > \lambda^{\text{max}} (\delta, \rho; \gamma) \right)$$

where $u$ is the number of groups.
\[ \lambda^{\max}(\delta, \rho) \] in the BCT analysis becomes \[ \lambda^{\max}(\delta, \rho; \gamma) \] for each \( \gamma := \frac{m}{n} \in [\rho, \delta^{-1}) \) by substituting \( \gamma \) for \( \rho \); \( \gamma = \rho \) recovers BCT

Larger values of \( m \) decrease the combinatorial term at the cost of increasing \( \lambda^{\max}(A^{*}A_{M}) \)

Interplay btw number & size of groups, \( \Rightarrow \) optimizing over \( \gamma \)

There exist an optimal \( \gamma \); shown below and proof trivial

Consequently, \( \lambda^{\max}(\delta, \rho) := \min_{\gamma} \lambda^{\max}(\delta, \rho; \gamma) \) and our upper bound is \( U^{BT}(\delta, \rho) = \lambda^{\max}(\delta, \rho) - 1 \)
Form groups $G_i := \{K\}$ for $K \subset \Omega := \{1, 2, \ldots, N\}$ and $M_i := \bigcup_{K \subset G_i, |K| = k} K$ with $|M_i| = m \geq k$.

Define $G := \bigcup_{i=1}^{u} G_i$ such that $|G| \geq \binom{N}{k}$, to have a covering.

**Lemma**

Set $r = \left(\frac{N}{k}\right)\left(\frac{m}{k}\right)^{-1}$ and draw $u := rN$ $M_i$ sets uniformly at random from the $\binom{N}{m}$ possible $M_i$ sets. With $G$ defined as above,

$$P \left[ |G| < \binom{N}{k} \right] < C\left(\frac{k}{N}\right)N^{-1/2}e^{-N(1-\ln 2)}, \text{ where } C(p) \leq \frac{5}{4} \left(2\pi p(1-p)\right)^{-1/2}$$

**Corollary**

Given the above lemma, as $n \to \infty$ in the proportional-growth asymptotics, the probability that all the $\binom{N}{k}$ $K \subset \Omega$ are covered by $G$ converges to one exponentially in $n$. 
Proof.

- Groups with \( m \geq k \) distinct elements, contains \( \binom{m}{k} K \subset \Omega \)
  \( \Rightarrow \) at least \( \binom{N}{k} \frac{m}{k}^{-1} =: r \) groups to cover each \( K \)

- For any random group, \( M_i \& K \), \( \mathbf{P}(M_i \supset K) = 1/r \) and 
  \[ \mathbf{P}(G \nsubseteq K) = (1 - 1/r)^u \leq \exp(-u/r) \]

- A union bound over \( \binom{N}{k} K \), yields 
  \[ \mathbf{P}[|G| < \binom{N}{k}] < \binom{N}{k} e^{-u/r} \]

- The RHS of Sterling’s Inequality gives 
  \[ \binom{N}{pN} \leq \frac{5}{4} (2\pi p(1 - p)N)^{-1/2} e^{NH(p)}, \quad H(p) \leq \ln 2 \text{ for } p \in [0, 1] \]

- Choosing \( u = rN \) completes proof of lemma.
- Letting \( n \to \infty \) proves the corollary.
Algorithms for calculating lower bounds of $L(k, n, N; A)$ & $U(k, n, N; A)$ by Dossal et. al. and Journée et. al. respectively
Sharpness ratios: 

Improvement ratios: 

<table>
<thead>
<tr>
<th>RIC bounds</th>
<th>C.T.</th>
<th>B.C.T.</th>
<th>B.T.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor of empirical data</td>
<td>2.74</td>
<td>1.83</td>
<td>1.57</td>
</tr>
</tbody>
</table>

- Factor 1.57 decreases to about 1.05 for $\rho < \frac{1}{100}$, where the CS results are applicable.
- BCT suffers from excessive overestimation when $\delta \rho \approx 1/2$
Phase Transitions based on BCT bounds \[ \text{[B.C.T.T., 2011]} \]
Small improvement on phase transitions, \( \approx 0.5 \sim 1\% \) higher

<table>
<thead>
<tr>
<th>Bounds</th>
<th>( \ell_1 )</th>
<th>IHT</th>
<th>SP</th>
<th>CoSaMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blanchard et. al.</td>
<td>317( k )</td>
<td>907( k )</td>
<td>3124( k )</td>
<td>4923( k )</td>
</tr>
<tr>
<td>Bah &amp; Tanner</td>
<td>314( k )</td>
<td>902( k )</td>
<td>3116( k )</td>
<td>4913( k )</td>
</tr>
</tbody>
</table>

Conditions giving the phase transitions are driven by \( L(\delta, \rho) \); precisely depending on \( 1 - L(\delta, \rho) \)

Our improvement has been greater in \( U(\delta, \rho) \) than in \( L(\delta, \rho) \)

But also improvement decreases from a max factor of 1.57 to about 1.05 in this regime of \( \delta \) and \( \rho \)
Our bounds valid for finite \((k, n, N)\), satisfying specified probs.

- Probabilities extremely small, even for small \((k, n, N)\) and \(\epsilon\)

<table>
<thead>
<tr>
<th>(k)</th>
<th>(n)</th>
<th>(N)</th>
<th>(\epsilon)</th>
<th>(Prob)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>200</td>
<td>2000</td>
<td>10(^{-3})</td>
<td>2.9 \times 10^{-2}</td>
</tr>
<tr>
<td>200</td>
<td>400</td>
<td>4000</td>
<td>10(^{-3})</td>
<td>9.5 \times 10^{-3}</td>
</tr>
<tr>
<td>400</td>
<td>800</td>
<td>8000</td>
<td>10(^{-3})</td>
<td>2.9 \times 10^{-3}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(k)</th>
<th>(n)</th>
<th>(N)</th>
<th>(\epsilon)</th>
<th>(Prob)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>200</td>
<td>2000</td>
<td>10(^{-5})</td>
<td>2.8 \times 10^{-18}</td>
</tr>
<tr>
<td>200</td>
<td>400</td>
<td>4000</td>
<td>10(^{-5})</td>
<td>9.1 \times 10^{-32}</td>
</tr>
<tr>
<td>400</td>
<td>800</td>
<td>8000</td>
<td>10(^{-5})</td>
<td>2.8 \times 10^{-58}</td>
</tr>
</tbody>
</table>
Theorem (Fixed $\delta$ and $\rho \to 0$)

Let $\tilde{U}(\delta, \rho)$ and $\tilde{L}(\delta, \rho)$ be the approximations of $U(\delta, \rho)$ and $L(\delta, \rho)$ respectively. For a fixed $\delta$ as $\rho \to 0$,

$$\tilde{L}(\delta, \rho) = \tilde{U}(\delta, \rho) = \sqrt{2\rho \log (\delta^{-2}\rho^{-3}) + 6\rho}$$

Theorem ($\rho$ as a function of $\delta$)

Let $\rho_\gamma(\delta) = \frac{1}{\gamma \log(\delta^{-1})}$ and let the approximations of $U(\delta, \rho)$ & $L(\delta, \rho)$ be $\tilde{U}(\delta, \rho_\gamma(\delta))$ & $\tilde{L}(\delta, \rho_\gamma(\delta))$ respectively. For a fixed $\gamma$ as $\delta \to 0$,

$$\tilde{U}(\delta, \rho_\gamma(\delta)) = \sqrt{2\rho \log (\delta^{-2}\rho^{-3}) + \frac{2}{3}\rho \log (\delta^{-2}\rho^{-3})}$$

$$\tilde{L}(\delta, \rho_\gamma(\delta)) = \sqrt{2\rho \log (\delta^{-2}\rho^{-3}) - \frac{2}{3}\rho \log (\delta^{-2}\rho^{-3})}$$
Corollary (From Theorem of $\rho$ as a function of $\delta$)

Let $\rho_\gamma(\delta) = \frac{1}{\gamma \log(\delta^{-1})}$ and let $\tilde{U}(\delta, \rho_\gamma(\delta))$ and $\tilde{L}(\delta, \rho_\gamma(\delta))$ be the approximations of $U(\delta, \rho)$ and $L(\delta, \rho)$ respectively. In the limit, $\delta \to 0$ and $\gamma \to \infty$, both $U(\delta, \rho_\gamma(\delta))$ and $L(\delta, \rho_\gamma(\delta))$ converge to $f(\gamma) := \sqrt{2}\gamma^{-1/2}$. 

\[ f(\gamma) = 2\sqrt{\gamma} \]

![Graph 1](image1)

![Graph 2](image2)
Corollary (Sampling Theorem for $\ell_1$, IHT, CoSaMP & SP)

Given a sensing matrix, $A$, of size $n \times N$ whose entries are drawn i.i.d. from $\mathcal{N}(0, 1/n)$, in the limit as $n/N \to 0$ the sufficient number of measurements for CS algorithms is $n \geq \gamma k \log(N/n)$, with

- $\gamma = 36$ for $\ell_1$-minimization,
- $\gamma = 93$ for Iterative Hard Thresholding (IHT),
- $\gamma = 272$ for Subspace Pursuit (SP) and
- $\gamma = 365$ for Compressed Sampling Matching Pursuit (CoSaMP).

Derivation uses approximations and recovery conditions for greedy algorithms $\mu^\text{alg}(\delta, \rho_\gamma(\delta)) = 1$ \cite{B.C.T.T., 2011}

$\gamma = 2e$ is known to be tight for $\ell_1$ \cite{Donoho & Tanner, 2009}
**Corollary (Sampling Theorem for OMP)**

*Given a sensing matrix, \( A \), of size \( n \times N \) whose entries are drawn i.i.d. from \( N(0, 1/n) \), in the limit as \( k/n \to 0 \) the sufficient number of measurements for Orthogonal Matching Pursuit (OMP) is*

\[
n \geq 16k^2 \log(N/k) + 8k^2 \log(n/k) + 24k^2.
\]

- OMP requires \( O\left(k^2 \log(N/k)\right) \) measurements to guarantee exact recovery  [Davenport & Wakin, 2010]
- Derivation uses approximation for fixed \( \delta \) with \( \rho \to 0 \) and recovery condition \( U(\delta, \rho) < \sqrt{k} - \sqrt{k - 1} \)  [Huang et. al., ’10]
For fast and efficient implementation sparse random matrices are preferable to dense ones.

But most sparse random matrices don’t satisfy RIP\(_2\); example: \{0, 1\} matrices with \(d\) ones per column do not satisfy RIP\(_2\), unless \(n = \Omega(k^2)\) [Chandar, 2007]

**Good News!** Adjacency matrices of unbalanced expander graphs satisfy RIP\(_1\) [Xu & Hassibi, ’07; Berinde et. al., ’08]

**Definition (RIP\(_1\) for Expanders)**

Let \(A \in \mathbb{R}^{n \times N}\) be the adjacency matrix of an unbalanced \((k, d, \epsilon)\)-expander graph \(G\), then for any \(k\)-sparse vector \(x \in \mathbb{R}^N\) we have:

\[
(1 - L_k) d \|x\|_1 \leq \|Ax\|_1 \leq d \|x\|_1
\]
Definition (Unbalanced Expander Graphs)

\( G = (U, V, E) \) is an unbalanced \((k, d, \epsilon)\)-expander if it is a bipartite graph with \(|U| = N\) left vertices, \(|V| = n\) right vertices and has a regular left degree \(d\), such that any \(X \subseteq U\) with \(|X| \leq k\) has \(|\Gamma(X)| \geq (1 - \epsilon) d|X|\) neighbors.

- Objects well-studied in theoretical computer science and coding theory but there are construction issues
- Berinde et. al. showed that \(A\) of a \((k, d, \epsilon/2)\)-expander satisfies RIP\(_1\) which means \(L_k = 2\epsilon\)
Theorem (Existence of Expanders)

Let $F$ be the event that a given random $G$ fails to be an unbalanced $(k, d, \epsilon)$-expander.

\[ P(F) \propto e^{n\Psi_{\epsilon,d}(\delta,\rho)} \]

- Therefore in p.g.a $P(F) \to 0$ if $\Psi_{\epsilon,d}(\delta,\rho) < 0$
- For any $\epsilon, d$, $\exists \rho_{\epsilon,d}^{XP}(\delta)$ such that $\Psi_{\epsilon,d}(\delta,\rho_{\epsilon,d}(\delta)) = 0$

Strong Phase Transitions

- The phase transition $\rho_{\epsilon,d}^{XP}(\delta)$ increases with $\epsilon$
- For each $\epsilon$, $\exists d$ optimal and an explicit expression for this $d$ is derived from $\Psi_{\epsilon,d}(\delta,\rho_{\epsilon,d}(\delta))$
Compressed Sensing (CS) and Related Areas
Restricted Isometry Property (2) - RIP_2
Asymptotic Approximation of RIC_2 Bounds
Conclusion

Sparse random matrices and RIP_1
Performance Guarantees for ℓ_1

ℓ_1 guarantees in sparse vs. dense

- ℓ_1 recovery is guaranteed if \( \epsilon \leq 1/6 \) which is equivalent to \( L_k \leq 1/3 \) [Berinde et. al., 2008]
- Phase transitions plots comparing ℓ_1 performance

- In the worst-case dense wins sparse but the average-case seem to tell a different story
The random matrix quantity, RIC is an important tool of analysis in CS and related areas.

Improvement is earlier bounds achieved by grouping of submatrices with significant column overlap.

Bounds clear improvements on prior bounds and are consistent with empirically observed data.

Asymptotic approximation of bounds lead to sampling theorems consistent with CS literature.

RIP₁ holds for (sparse) adjacency matrices of expanders.

ℓ₁ performance guarantees better for dense than sparse.
Note:

In the spirit of reproducible research, software and web forms that evaluate $L^{BT}(\delta, \rho)$ and $U^{BT}(\delta, \rho)$ are publicly available at http://ecos.maths.ed.ac.uk/ric_bounds.shtml

References:


THANK YOU