RECONSTRUCTION OF SPARSE SIGNALS FROM $\ell_1$ DIMENSIONALITY-REDUCED
CAUCHY RANDOM-PROJECTIONS

Gonzalo R. Arce, Daniel Otero, Ana B. Ramirez, and Jose-Luis Paredes

†Department of Electrical and Computer Engineering
University of Delaware, Newark, DE, USA, 19716
email: arce@udel.edu

ABSTRACT

Dimensionality reduction via linear random projections are used in numerous applications including data streaming, information retrieval, data mining, and compressive sensing (CS). While CS has traditionally relied on normal random projections, corresponding to $\ell_2$ distance preservation, a large body of work has emerged for applications where $\ell_1$ approximate distances may be preferred. Dimensionality reduction in $\ell_1$ use Cauchy random projections that multiply the original data matrix $B \in \mathbb{R}^{D \times n}$ with a Cauchy random matrix $R \in \mathbb{R}^{n \times k}$, resulting in a projected matrix $C \in \mathbb{R}^{D \times k}$. This paper focuses on developing signal reconstruction algorithms from Cauchy random projections, where the large suite of reconstruction algorithms developed in compressive sensing perform poorly due to the lack of finite second-order statistics in the projections. In particular, a set of regularized coordinate-descent Myriad regression based reconstruction algorithms are developed using, both $l_0$ and Lorentzian norms as sparsity inducing terms. The $l_0$-regularized algorithm shows superior performance to other standard approaches. Simulations illustrate and compare accuracy of reconstruction.

Index Terms— Dimensionality reduction, compressed sensing, Cauchy random projections, sketching.

1. INTRODUCTION

Dimensionality reduction methods by linear random projections enable the mapping of a set of high-dimensional data points into a set of points in low-dimension, such that both sets have similar distance properties. Linear random projections multiply the original data matrix $B \in \mathbb{R}^{D \times n}$ with a random matrix $R \in \mathbb{R}^{n \times k}$, resulting in a projected matrix $C \in \mathbb{R}^{D \times k}$. If $k \ll \min(n,D)$, then it is much more efficient to compute certain summary statistics, such as pairwise distances, from $C$ than from the original data set $B$. Moreover, $C$ may be small enough to store in physical memory while $B$ is often too large. The choice of the random projection matrix $R$ depends on which norm is preferred. Indyk [1] proposed constructing $R$ from i.i.d. samples of $p$-stable distributions, for dimension reduction in $l_p$ ($0 < p \leq 2$). In the stable distribution family [2], normal is $2$-stable and Cauchy is $1$-stable. In normal random projections [3], one can estimate the original pairwise $l_2$ distances of $B$ directly using the corresponding $l_2$ distances of $C$, and the Johnson-Lindenstrauss (JL) lemma [4] provides the performance guarantee. In Cauchy random projections, the situation is more involved as the first moments $E(\sum ||c_i - c_j||)$ are undefined. Nonetheless, Indyk [1] has shown a weaker analog of the Johnson-Lindenstrauss bound for the $l_1$ norm for stable random projections. The interest in Cauchy random projections arise since the $l_1$ norm is more robust to noise, missing data, and outliers, than the $l_2$ norm.

Given the original data points $B$, let $\{b_i^T\}_{i=1}^D \in \mathbb{R}^n$ be the $i^{th}$ row of $B$. Further, denote $\{r_{i,j}\}_{i=1}^D \in \mathbb{R}^k$ as the entries of the Cauchy random projection matrix $R \in \mathbb{R}^{n \times k}$. The $i^{th}$ row of $C$, $\{c_i^T\}_{i=1}^D \in \mathbb{R}^k$, referred to as a Cauchy random variable with the scale parameter being the leading two rows, $c_1$, and $c_2$, in $B$, and the leading two rows, $c_1$ and $c_2$, in $C$. Define $u_j = (c_{1,j} - c_{2,j}) = \sum_{i=1}^n r_{i,j}(b_{1,i} - b_{2,i})$ for $j = 1, 2, \cdots, k$. Then, if $r_{i,j}$ i.i.d. samples from the standard Cauchy distribution, i.e., $r_{i,j} \sim C(0,1)$, then by the $1$-stability property of Cauchy [2] it follows that

$$u_j = c_{1,j} - c_{2,j} \sim C\left(0, \frac{\sum_{i=1}^n |b_{1,i} - b_{2,i}|}{k}\right),$$

where the Cauchy random variable $z \sim C(0,\gamma)$ has the density $f(z) = \frac{\gamma}{\pi (\gamma^2 + z^2)}$, $-\infty < z < \infty$, where $\gamma > 0$ is the scale parameter. Thus, the projected differences $u_j = (c_{1,j} - c_{2,j})$ are also Cauchy random variables with the scale parameter being the $l_1$ distance, $d = \|b_1 - b_2\| = \sum_{i=1}^n |b_{1,i} - b_{2,i}|$, in the original space. Therefore, in Cauchy random projections, the $l_1$ distance estimation problem reduces to estimating the Cauchy scale parameter of $C(0,d)$ from $k$ i.i.d. samples $u_j \sim C(0,d)$. The work in [5] elaborates on various nonlinear estimates of $d$.

This paper focuses on extending the above theory of dimensionality reduction with the $l_1$ norm, to allow signal reconstruction. That is, given the Cauchy random projected observations $C$, we not only seek to estimate $l_1$ distances of the original data but we also seek to reconstruct it with negligible error. Reconstruction is desirable in applications where the true data $B$ may be needed for inspection and verification. Reconstruction is also required in the broad class of compressive sampling (CS) applications. Note that CS is a special case of dimensionality reduction by linear projections and consequently the reconstruction problem at hand is indeed similar to that in the compressive sensing literature. The difference, however, lies in that Cauchy random projections are used here, rather than random projections that are either normally distributed or have finite second-order statistics. Since Cauchy random projections have undefined second-order statistics, the large suite of reconstruction methods developed for compressive sensing fail. The methods developed here build on a rich class of robust regression algorithms recently developed for signal processing under stable models [2]. In particular, the reconstruction problem is formulated as an iterative coordinate-descent parameter estimation problem as described next.
2. RECONSTRUCTION FROM CAUCHY PROJECTIONS: PROBLEM FORMULATION

Given the set of sketches $C \in \mathbb{R}^{D \times k}$, we seek the reconstruction of any particular row of the original sparse data matrix $B \in \mathbb{R}^{D \times n}$. Without loss of generality, consider the reconstruction of the $t^{th}$ row $b_t = [b_{t,1}, b_{t,2}, \ldots, b_{t,n}]$ of $B$, from its sketch $c_t = [c_{t,1}, c_{t,2}, \ldots, c_{t,k}]$. A common criterion widely used in the compressive sensing literature is to reconstruct the sparse signal by minimizing the norm of a residual error subject to the constraint that the signal is sparse [6]. In the context of Cauchy projections, a suitable formulation of this approach is

$$b_t = \arg \min_{b_t} \| c_t - R^T b_t \|_{ll} + \tau \| b_t \|_0 \tag{2}$$

where $\|Z\|_{ll}$ denotes the Lorentzian $L_2$-norm $^1$ of a vector $Z$ defined as $\sum_{i} \log (K^2 + Z^2)$ for $K > 0$, and where $\| b \|_0$ denotes the $l_0$ quasi-norm that outputs the number of nonzero components in its argument, and where $\tau$ is a regularization parameter that balances the conflicting goals of minimizing the norm of the residual errors while yielding, at the same time, a sparse solution on $b_t$. The Lorentzian error norm is characterized by a re-descendent influence function where the influence of a large outlying sample on the Lorentzian norm increases up to a point, after which it starts to decrease (redescend) as the error grows. Thus, the Lorentzian norm is more robust than the $l_1$ and $l_2$ norms and, in fact, it has optimality properties for Cauchy distributed samples [2].

Unfortunately, the optimization problem in (2) is $NP$-hard whose direct solution even for modest-sized signal is unfeasible. A first approach to reduce the complexity involved in (2) is to approximate the $l_0$ norm by a convex norm that at the same time induces sparsity. This convex-relaxation approach has been widely used as a powerful optimization tool to solve multivariable optimization problems. More recently, it has been proposed for solving regularized linear regression problems in [8]. The key to the success of this algorithm relies in the fact that it is much easier to solve a one-dimensional minimization problem than a multidimensional one. Furthermore, if a closed form expression is available for the solution of the single-variable minimization subproblem, the multidimensional optimization problem reduces to an update of each signal component in an optimal fashion. To this end, a closed form expression for the solution of the one-dimensional $l_0$-regularized $L_2$-optimization problem is derived and compared to the one obtained for the solution of the norm relaxation approach, where the $l_0$ norm in (2) is replaced by the Lorentzian norm.

$^1$The $l_2$ norm is generally used in the CS literature since Gaussian random projections are typically used in CS. The $l_2$ norm, however, is not applicable here as Cauchy projections lack second-order statistics. The $l_p$-norm of a vector $Z$ defined as $\left(\sum_i \|Z_i\|^p\right)^{1/p}$ for $p < 1$, is also suitable in (2) [2].

3. REGULARIZED ITERATIVE COORDINATE-DESCENT RECONSTRUCTION

For notational simplicity we drop the index $t$ as we focus on the reconstruction of the $t^{th}$ row hereafter. The reconstruction of any other row then follows a similar procedure. The elements of the $t^{th}$ sketch can thus be written as $c_{t,i} = \sum_{j=1}^{k} r_{i,j} b_j$, for $i = 1, 2, \ldots, k$, where $r_{i,j} \sim \mathcal{C}(0, 1)$. Assuming the value of $r_{i,j}$ is known, or can be re-generated through pseudorandom noise generation, the sketch elements can be scaled into the observations

$$c_{t,i} = \frac{r_i}{r_{i,t}} b_t + \frac{\sum_{j=1, j \neq t}^{k} r_{i,j} b_j}{r_{i,t}}, \quad i = 1, \ldots, k. \tag{3}$$

The above can be thought of as a classical (deterministic) location parameter estimation problem $Z_{i,t} = \beta + \eta_{i,t}$, where the observation samples are $Z_{i,t} = \hat{c}_{t,i}$, the unknown location parameter is $\beta = b_t$, and the noise samples are $\eta_{i,t} = \frac{\sum_{j=1, j \neq t}^{k} r_{i,j} b_j}{r_{i,t}}$. Since the noise term is a weighted combination of $(n-1)$ i.i.d. standard Cauchy random variables, and since $r_{i,t}$ is assumed known, then $\eta_{i,t} \sim \mathcal{C}(0, \gamma_{i,t})$ where $\gamma_{i,t} = \frac{\| b \|_1}{r_{i,t}} = \sum_{i=1, i \neq t}^{n} |b_i|$. Thus, given the observations $Z_{i,t}$, each obeying the Cauchy distribution with a common location parameter $\beta$ but varying scaling factor $\gamma_{i,t}$, the maximum likelihood estimate of $\beta$ is

$$\hat{\beta} = \arg \min_{\beta} P(\beta) \triangleq \arg \min_{\beta} \frac{1}{k} \left[ 1 + \frac{1}{\gamma_{i,t}} \| Z_{i,t} - \beta \|_2^2 \right] \tag{4}$$

Alternatively, we can write (4) in the form of a Myriad filter estimate

$$\hat{\beta} = \arg \min_{\beta} Q(\beta) \triangleq \arg \min_{\beta} \frac{1}{k} \sum_{i=1}^{k} \log \left[ K^2 + W_{i,t} \sum_{i=1}^{k} \| Z_{i,t} - \beta \|_2^2 \right]$$

$$= \text{MYPRIAD} \{K, W_{i,t} \circ Z_{i,t}, \ldots, W_{t,t} \circ Z_{t,t}\}, \tag{5}$$

where $\gamma_{i,t} \triangleq \frac{K}{\sqrt{W_{i,t}}} > 0$, and $W_{i,t} \circ Z_{i,t}$ represents the weighting operation in (5) [2, 9]. Thus $\hat{\beta}$ is the global minimizer of $P(\beta)$ as well as of $Q(\beta) \triangleq \log (P(\beta))$. Note that if the scale $\gamma_{i,t} \to \infty$ (or $r_{i,t} \to 0$), $Z_{i,t}$ becomes unreliable and consequently this term drops out of $P(\beta)$ and $Q(\beta)$ and is thus effectively ignored in the estimate. In the limit as $K \to \infty$, with the weights $\{W_{i,t}\}$ held constant, it can be shown that $Q(\beta)$ exhibits a single local extremum and the weighted myriad reduces to the normalized linear estimate [9]. At the other extreme, as $K \to 0$ the weighted myriad estimator adopts a mode-like behavior assigning more credibility to samples in a cluster with the largest plurality and, in fact, the estimator becomes a selection-type where the output will always be one of the samples in the cluster [9]. This property will play a key role later on as we include sparse-inducing regularization into (5). See [2, 9] for further properties of the Myriad.

The location estimation problem in (3) assumed that the $b_t$ for $i = 1, 2, \ldots, n; i \neq t$ contained in the noise term are known. Since this is not the case initially, the noise term statistics need to be estimated from the initial observations $Z_{i,t}^{(0)} = \frac{c_{t,i}}{r_{i,t}}$. In particular, the unbiased estimate of dispersion for the noise term, described in [5], is used as

$$\gamma_{i,t}^{(0)} = \cos^k \left( \frac{\pi}{2K} \right) \prod_{i=1}^{k} \left( Z_{i,t}^{(0)} - T_m(Z_{i,t}^{(0)}) \right), \quad i = 1, \ldots, n, \tag{6}$$

where $T_m(Z_{i,t}^{(0)})$ is a trimmed-mean estimate of location [2]. Given the initial observations $Z_{i,t}^{(0)}$ and the corresponding $\gamma_{i,t}^{(0)}$, the initial signal estimates $b_{t}^{(1)}$, for each coordinate, are then found with (4).
After this first iteration, the reconstruction algorithm follows an iterative coordinate-descent solution approach where we hold constant all but one of the entries of \( B(p) \), we then estimate the entry that is allowed to vary, and then we move on to estimate the entry in the next coordinate. This procedure is repeated until a performance criterion is achieved. Thus, each entry of the sparse vector is iteratively estimated based on previous estimated values of the other entries. It is important to note, however, that after the first iteration is completed the location estimation problem for \( p \geq 2 \) is redefined as

\[
\hat{c}_i - \sum_{j=1, j \neq i}^{n} r_{i,j} b_{j}(0) = b_{i} + \sum_{j=1, j \neq i}^{n} r_{i,j} b_{j}(p),
\]

where the \( j \)-th signal component \( b_{j} \) is represented as a sum of an estimated term plus an unknown residual, i.e., \( b_{j} = b_{j}^{(0)} + b_{j}^{(p)} \).

The formulation in (7) remains a location estimation problem \( Z_{0,i} = \beta + \eta_{0,i}^{(p)} \) where the observation samples \( Z_{i}^{(p)} \) are now the residual terms in the left of (7), and the noise term \( \eta_{0,i}^{(p)} \) is still Cauchy distributed with dispersion \( \gamma_{0,i}^{(p)} \). The dispersion at the \( p \)-th iteration is thus estimated with the geometric mean given in (6).

A. \( l_{0} \)-Regularized Coordinate Descent Reconstruction

The Myriad coordinate estimation in (5) is optimal but will not induce a sparse signal reconstruction. The \( l_{0} \)-regularized reformulation of (5) that favors sparsity is given by

\[
\hat{\beta}_0 = \arg\min_{\beta} Q_{0}(\beta) \triangleq \arg\min_{\beta} Q(\beta) + \tau \| \beta \|_0
\]

where the regularisation term above considers two cases: \( \beta = 0 \) and \( \beta \neq 0 \) as follows

\[
Q_{0}(\beta) = \left\{ \begin{array}{ll}
Q(0) & \text{if } \beta = 0 \\
Q(\beta) + \tau & \text{if } \beta \neq 0,
\end{array} \right.
\]

thus the \( l_{0} - LL \) regularized estimate of \( \beta_{0} \) reduces to

\[
\hat{\beta}_0 = \left\{ \begin{array}{ll}
0 & \text{if } Q(0) < Q(\beta) + \tau \\
\arg\min_{\beta} Q(\beta) & \text{otherwise.}
\end{array} \right.
\]

B. Norm Relaxation Coordinate Descent Reconstruction \( (LL - LL) \)

A second approach to induce sparsity in (5) is to approximate the \( l_{0} \)-norm by a norm that is mathematically more tractable. Our approach to convex relaxation exploits the logarithm form of the Lorentzian cost function to define the following norm relaxation function

\[
\hat{\beta}_{LL} = \arg\min_{\beta} \sum_{i=1}^{k} \log \left[ K^2 + W_{0,i}(Z_{i,0} - \beta)^2 \right] + \log \left[ K^2 + W_{0,i}(Z_{0,i} - \beta)^2 \right]
\]

\[
= \text{MYRIAD} \left\{ K; W_{0,i} \circ Z_{0,i}, \cdots, W_{k,i} \circ Z_{k,i} \right\},
\]

where we let \( Z_{0,i} = 0 \) and where \( W_{0,i} \) is the weight assigned to the zero-valued sample. The additional zero-valued sample fed into the estimate will induce sparsity, provided that the corresponding weight \( W_{0,i} \) is sufficiently large.

To implement both iterative reconstruction algorithms presented above, it is critical to properly estimate the regularization parameters \( \tau \) and \( W_{0,i} \), as these parameters govern the sparsity of the solution. Consider first the design of the regularization weight \( W_{0,i} \) in (11) where it can be seen that sparsity will be induced if the weight \( W_{0,i} \) is large enough to pull the estimate towards \( Z_{0,i} = 0 \). From the mode-property of the weighted Myriad [9, 2], this will occur if \( \gamma_{0,i}^{2} \leq \min \{ Z_{0,i}^2 + \gamma_{0,i}^2 \} \).  

if \( W_{0,i} > (\min \{ Z_{0,i}^2 + \frac{1}{W_{0,i}} \})^{-1} \). The design of \( \tau \) in (10) is determined similarly. Since \( W_{0,i} \) is the regularization parameter affecting \( Z_{0,i} = 0 \) inside a logarithm term in (11), and since \( \tau \) affects \( Z_{0,i} = 0 \) directly, it follows that \( \tau \geq \log(W_{0,i}) \). Furthermore, our algorithm gradually reduces the value of the regularization parameter with each iteration by defining \( \tau = \tau_{0} p^{2} \), where \( p \) refers to the iteration number, and \( 0 < p < 1 \). The algorithm continues until \( \tau \) reaches a final value or until a total number of iterations is reached. The \( l_{0} \)-Regularized Coordinate Descent algorithm for sparse signal reconstruction is shown in Table 1. Note in step A, that the most recent estimated value for each entry is used for the estimation of subsequent entries in the same iteration step.

The following norm relaxation function

\[
\tau(0) = \log(10 \min \{ Z_{0,i}^{2} + \frac{1}{W_{0,i}} \})^{-1}
\]

Initial Estimated Signal \( b_{0} = 0 \)

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Step A</th>
<th>For ( \ell = 1, 2, \ldots, n ) compute</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{b}<em>{\ell} = \text{MYRIAD} \left{ W</em>{0,i} \circ Z_{0,i}, \cdots, W_{k,i} \circ Z_{k,i} \right} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \alpha(\theta) = \sum_{i=1}^{k} \log \left[ K^2 + W_{0,i}(Z_{0,i} - \theta)^2 \right] )</td>
<td></td>
</tr>
</tbody>
</table>
|           | \( \hat{b}_{\ell}^{(p)} = \left\{ \begin{array}{ll}
0 & \text{if } \alpha(0) - \alpha(b_{\ell}) < \tau(0) \\
\hat{b}_{\ell} & \text{otherwise,}
\end{array} \right. \) | |
| Step B    | \( \tau^{(p)} = \tau_{0} p^{2} \) | |
|           | \( \hat{b}_{\ell}^{(p+1)} = b_{\ell}^{(p)} \) | |
|           | \( \gamma_{(p+1),i} = \cos(k) \prod_{i=1}^{P} \left( Z_{0,i}^{(p)} - T_{m}(Z_{i,0}^{(p)}) \right) \) | |
| Step C    | If \( E^{(p)} = |e - R^{(p)}| / |e| < E_{0} \), end; else go to Step A. | |
| Output    | Recovered sparse signal \( b^{(p)} \) | |

4. COMPUTER SIMULATIONS AND PERFORMANCE

The performance of the proposed \( l_{0} - LL \) algorithm is evaluated in the recovery of a sparse signal from a sketch with additive Gaussian noise of various levels. Comparisons with the \( l_{1} \)-regularized least squares (\( l_{1} \)-ls) and the Orthogonal Matching Pursuit algorithms are included [6]. In addition, some examples of reconstruction of signals from noiseless sketches of the proposed \( l_{0} - LL \) and \( LL - LL \) algorithms are presented and compared with the \( l_{1} \)-regularized least squares (\( l_{1} \)-ls) algorithm. In all the simulations, the \( n \)-sample sparse signal is generated by randomly placing the location of the non-zero entries which are generated by a uniform random distribution and with amplitudes in the interval \((-1, 1)\). The projection matrix is generated with i.i.d. draws of a standard Cauchy distribution. For each random trial, a new Cauchy projection matrix, a new sparse signal and a random Gaussian noise realization are generated.
Mean Square Error (dB): The reconstruction capability of the algorithm is tested in the recovery of a sparse signal by finding the MSE obtained when a stopping criterion is reached. Thus, the iterative algorithm is ended as soon as $E^{(p)} = ||c - Rh||^2 < 10^{-3}$ is reached. Since Cauchy projections have infinite-variance, the input SNR becomes less meaningful and thus we use the Geometric Signal-to-Noise Ratio (G-SNR) defined as \[ G - SNR = \frac{S_0}{S_0 + N_0} \]
where $S_0$ is the geometric power of the projected signal and $N_0$ is the geometric power of the additive Gaussian noise. Figure 1(a) shows the MSE (dB) obtained for different levels of noise in G-SNR(dB) for the proposed $l_0$ - LL method as well as that obtained for the $l_1$ - regularized least squares ($l_1$-ls) and the Orthogonal Matching Pursuit algorithms. In can be noticed in Fig 1(a) that for a GSNR of 20dB the proposed reconstruction $l_0$ - LL method achieves improvements in the MSE of up to 25dB over the L1-LS algorithm and of up to 32dB with respect to the OMP algorithm.

Frequency of exact reconstruction: Figure 1(b) depicts the probability of exact reconstruction for the various algorithms (averaged over 500 trials) as a function of the number of projections $k$. Exact reconstruction is assumed when the Mean Square Error obtained in $b$ is less than $10^{-3}$. The number of non-zero entries in the sparse signal is set to $S = 0.02n$, where $n$ is the dimension of the signal and the number of random projections $k$ gradually increases from 0.02$n$ to $n$. For this simulation, $\rho$ is set to 0.7 and $\tau_0$ the initial regularization parameter is set to be the logarithm of ten times the corresponding minimum weight $W_{0,c}$. Figure 1(b) illustrates that the $l_0$ - LL algorithm requires significantly less fewer measurements than the $l_1$ - ls and OMP algorithms to achieve the same probability of successful reconstruction.

Example of a signal reconstruction: Figure 2(a) illustrates the markedly differences between the projections of a sparse signal (top) with Gaussian random projections (bottom solid) and with Cauchy random projections (bottom dotted). The lack of finite 2nd-order statistics in the Cauchy projections is evident. Figure 2(b) depicts the signal reconstruction of the $l_1$ - ls algorithm, while 2(c) and 2(d) depicts the reconstruction of the $LL$ - $LL$ and $l_0$ - $LL$ algorithms when no measurement noise is present, respectively. Note that the $l_0$ - $LL$ reconstruction is superior. Results using $LL$ - $LL$ reconstruction algorithm can be improved if debiasing is used.

Fig. 2. Reconstruction of a sparse signal from a noiseless sketch $(n = 400, k = 100)$. (a) Top: Sparse signal. Bottom: Gaussian projections and Cauchy Projections. (b) $l_1$ - ls reconstruction. (c) $LL$ - $LL$ reconstruction. (d) $l_0$ - $LL$ reconstruction. Original signal: blue circles. Reconstruction: red signal.

5. CONCLUSION

We present an algorithm to reconstruct a sparse signal from dimensionality reduced sketches in $l_1$ that are obtained using Cauchy random projections. The proposed approach recasts the reconstruction problem as a location parameter estimation where each entry of the sparse signal is iteratively estimated by a $l_0$ regularized Myriad operator. This algorithm follows a coordinate descent approach for solving the multidimensional $l_0$ - $LL$ regression problem. The proposed methods extend the theory of dimensionality reduction with the $l_1$ norm, to allow signal reconstruction from Cauchy random projected sketches.

6. REFERENCES