Supplementary Material: A Simple Homotopy Algorithm for Compressive Sensing

Lijun Zhang[*] **Tianbao Yang**[†] **Rong Jin**[‡] **Zhi-Hua Zhou**[*]
[*]National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
[†]Department of Computer Science, the University of Iowa, Iowa City, USA
[‡]Department of Computer Science and Engineering, Michigan State University, East Lansing, USA
[‡]Institute of Data Science and Technologies at Alibaba Group, Seattle, USA
{zhanglj, zhouzh}@lamda.nju.edu.cn tianbao-yang@uiowa.edu rongjin@cse.msu.edu

A Proof of Lemma 2

The analysis is the same as that for Lemma 9.2 of Koltchinskii (2011), we include it for completeness. For any $\mathbf{x}, \mathbf{x}' \in \mathcal{K}_{d,s}$, we can always find two vectors $\mathbf{y}$, $\mathbf{y}'$ such that

$$\mathbf{x} - \mathbf{x}' = \mathbf{y} - \mathbf{y}', \ \|\mathbf{y}\|_0 \leq s, \ \|\mathbf{y}'\|_0 \leq s, \ \mathbf{y}^\top \mathbf{y}' = 0.$$

Thus

$$\begin{aligned}
&\langle \mathbf{x} - \mathbf{x}', UU^\top \mathbf{z} \rangle = \langle \mathbf{y}, UU^\top \mathbf{z} \rangle + \langle -\mathbf{y}', UU^\top \mathbf{z} \rangle \\
=&\|\mathbf{y}\|_2 \left\langle \frac{\mathbf{y}}{\|\mathbf{y}\|_2}, UU^\top \mathbf{z} \right\rangle + \|\mathbf{y}'\|_2 \left\langle \frac{-\mathbf{y}'}{\|\mathbf{y}'\|_2}, UU^\top \mathbf{z} \right\rangle \\
\leq&(\|\mathbf{y}\|_2 + \|\mathbf{y}'\|_2)\mathcal{E}_s(\mathbf{z}) \leq \mathcal{E}_s(\mathbf{z})\sqrt{2}\sqrt{\|\mathbf{y}\|_2^2 + \|\mathbf{y}'\|_2^2} \\
=&\mathcal{E}_s(\mathbf{z})\sqrt{2}\|\mathbf{y} - \mathbf{y}'\|_2 = \mathcal{E}_s(\mathbf{z})\sqrt{2}\|\mathbf{x} - \mathbf{x}'\|_2.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\mathcal{E}_s(\mathbf{z}) = \max_{\mathbf{w} \in \mathcal{K}_{d,s}} \mathbf{w}^\top UU^\top \mathbf{z} \\
\leq&\mathcal{E}_s(\mathbf{z}, \epsilon) + \sup_{\mathbf{x} \in \mathcal{K}_{d,s}, \mathbf{x}' \in \mathcal{K}_{d,s}(\epsilon), \|\mathbf{x} - \mathbf{x}'\|_2 \leq \epsilon} \langle \mathbf{x} - \mathbf{x}', UU^\top \mathbf{z} \rangle \\
\leq&\mathcal{E}_s(\mathbf{z}, \epsilon) + \sqrt{2}\epsilon\mathcal{E}_s(\mathbf{z})
\end{aligned}$$

which implies

$$\mathcal{E}_s(\mathbf{z}) \leq \frac{\mathcal{E}_s(\mathbf{z}, \epsilon)}{1 - \sqrt{2}\epsilon}.$$

B Proof of Lemma 3

Since

$$\begin{aligned}
&|\mathbf{w}^\top UU^\top \mathbf{z} - \mathbf{w}^\top \mathbf{z}| \\
=&\|\mathbf{w}\|_2 \left| \frac{1}{\|\mathbf{w}\|_2}\mathbf{w}^\top UU^\top \mathbf{z} - \frac{1}{\|\mathbf{w}\|_2}\mathbf{w}^\top \mathbf{z} \right| \\
\leq&\left| \frac{1}{\|\mathbf{w}\|_2}\mathbf{w}^\top UU^\top \mathbf{z} - \frac{1}{\|\mathbf{w}\|_2}\mathbf{w}^\top \mathbf{z} \right|,
\end{aligned}$$

without loss of generality, we can assume $\|\mathbf{w}\|_2 = 1$.

We decompose $\mathbf{z}$ as $\mathbf{z} = \mathbf{z}_\| + \mathbf{z}_\perp$, where

$$\mathbf{z}_\| = (\mathbf{z}^\top \mathbf{w})\mathbf{w}, \ \mathbf{z}_\perp = \mathbf{z} - \mathbf{z}_\|.$$

As a result

$$\begin{aligned}
\mathbf{w}^\top UU^\top \mathbf{z} =&\mathbf{w}^\top UU^\top \mathbf{z}_\| + \mathbf{w}^\top UU^\top \mathbf{z}_\perp \\
=&(\mathbf{z}^\top \mathbf{w})\|U^\top \mathbf{w}\|_2^2 + \mathbf{w}^\top UU^\top \mathbf{z}_\perp.
\end{aligned} \tag{5}$$

We first consider bounding $\|U^\top \mathbf{w}\|_2^2$. Notice that $U = \frac{1}{\sqrt{m}}[\mathbf{u}_1, \ldots, \mathbf{u}_m] \in \mathbb{R}^{d \times m}$, and we assume $\mathbf{u}_i$'s are independent, isotropic, and sub-Gaussian vectors. Then, for any fixed vector $\mathbf{x}$, with a probability at least $1 - e^{-C_1 m \epsilon^2}$, we have (Mendelson et al., 2008, Section 3.1)

$$(1 - \epsilon)\|\mathbf{x}\|_2^2 \leq \|U^\top \mathbf{x}\|_2^2 \leq (1 + \epsilon)\|\mathbf{x}\|_2^2$$

where $C_1 > 0$ is some constant. And thus, with a probability at least with a least $1 - e^{-\tau}$, we have

$$1 - C_1\sqrt{\frac{\tau}{m}} \leq \|U^\top \mathbf{w}\|_2^2 \leq 1 + C_1\sqrt{\frac{\tau}{m}} \tag{6}$$

for some constant $C_1 > 0$.

Next, we consider bounding $\mathbf{w}^\top UU^\top \mathbf{z}_\perp = \frac{1}{m}\sum_{i=1}^m \mathbf{w}^\top \mathbf{u}_i \mathbf{u}_i^\top \mathbf{z}_\perp$. Since $\mathbf{u}_i$'s are isotropic, we have

$$\mathrm{E}[\mathbf{w}^\top \mathbf{u}_i \mathbf{u}_i^\top \mathbf{z}_\perp] = \mathbf{w}^\top \mathbf{z}_\perp = 0.$$

Based on the property $\|\eta_1\eta_2\|_{\psi_1} \leq \|\eta_1\|_{\psi_2}\|\eta_2\|_{\psi_2}$ (Koltchinskii, 2009, Page 815), we know that $\mathbf{w}^\top \mathbf{u}_i \mathbf{u}_i^\top \mathbf{z}_\perp$ is a sub-exponential random variable, and

$$\begin{aligned}
\|\mathbf{w}^\top \mathbf{u}_i \mathbf{u}_i^\top \mathbf{z}_\perp\|_{\psi_1} \leq&\|\mathbf{u}_i^\top \mathbf{w}\|_{\psi_2}\|\mathbf{u}_i^\top \mathbf{z}_\perp\|_{\psi_2} \\
\leq&\|\mathbf{w}\|_2\|\mathbf{z}_\perp\|_2 \leq \|\mathbf{z}\|_2.
\end{aligned}$$

And thus $\{\mathbf{w}^\top \mathbf{u}_i \mathbf{u}_i^\top \mathbf{z}_\perp\}_{i=1}^m$ is a set of independent centered sub-exponential random variables. Following the Bernstein-type inequality for sub-exponential random variables (Vershynin, 2012, Proposition 5.16), with a probability at least $1 - e^{-\tau}$, we have

$$\left|\mathbf{w}^\top U U^\top \mathbf{z}_\perp\right| \leq C_2 \|\mathbf{z}\|_2 \sqrt{\frac{\tau}{m}} \tag{7}$$

for some constant $C_2 > 0$.

Putting everything together, with a probability at least $1 - 2e^{-\tau}$, we have

$$\begin{aligned}
&\left|\mathbf{w}^\top U U^\top \mathbf{z} - \mathbf{w}^\top \mathbf{z}\right| \\
\overset{(5)}{=} &\left|(\mathbf{z}^\top \mathbf{w})\|U^\top \mathbf{w}\|_2^2 + \mathbf{w}^\top U U^\top \mathbf{z}_\perp - \mathbf{w}^\top \mathbf{z}\right| \\
\leq &\left|\mathbf{w}^\top \mathbf{z}\right| \left|\|U^\top \mathbf{w}\|_2^2 - 1\right| + \left|\mathbf{w}^\top U U^\top \mathbf{z}_\perp\right| \\
\overset{(6), (7)}{\leq} & C_1 \sqrt{\frac{\tau}{m}} \left|\mathbf{w}^\top \mathbf{z}\right| + C_2 \|\mathbf{z}\|_2 \sqrt{\frac{\tau}{m}} \\
\leq &(C_1 + C_2) \|\mathbf{z}\|_2 \sqrt{\frac{\tau}{m}}.
\end{aligned}$$

C Proof of Theorem 2

Let $\mathcal{X}$ and $\mathcal{Y}$ be the support set of $\mathbf{x}$ and $\mathbf{y}$, respectively. If $|\mathcal{X}| \leq s$, we have

$$\|\mathbf{x}^s - \mathbf{y}\|_2 = \|\mathbf{x} - \mathbf{y}\|_2.$$

Thus, in the following, we only need to consider the case $|\mathcal{X}| > s$.

Let $\mathcal{A}$ be the indices of the s largest elements in $\mathbf{x}$, and $\mathcal{B} = \mathcal{X} \setminus \mathcal{A}$. Then, we have

$$\begin{aligned}
\|\mathbf{x} - \mathbf{y}\|_2^2 = &\sum_{i \in \mathcal{A} \setminus \mathcal{Y}} x_i^2 + \sum_{i \in \mathcal{A} \cap \mathcal{Y}} (x_i - y_i)^2 \\
&+ \sum_{i \in \mathcal{B} \cap \mathcal{Y}} (x_i - y_i)^2 + \sum_{i \in \mathcal{B} \setminus \mathcal{Y}} x_i^2, \\
\|\mathbf{x}^s - \mathbf{y}\|_2^2 = &\sum_{i \in \mathcal{A} \setminus \mathcal{Y}} x_i^2 + \sum_{i \in \mathcal{A} \cap \mathcal{Y}} (x_i - y_i)^2 + \sum_{i \in \mathcal{B} \cap \mathcal{Y}} y_i^2.
\end{aligned}$$

Since

$$|\mathcal{A} \setminus \mathcal{Y}| + |\mathcal{A} \cap \mathcal{Y}| = |\mathcal{A}| = s \geq |\mathcal{Y}| = |\mathcal{A} \cap \mathcal{Y}| + |\mathcal{B} \cap \mathcal{Y}|$$

we have $|\mathcal{A} \setminus \mathcal{Y}| \geq |\mathcal{B} \cap \mathcal{Y}|$. As a result, we must have

$$\sum_{i \in \mathcal{B} \cap \mathcal{Y}} x_i^2 \leq \sum_{i \in \mathcal{A} \setminus \mathcal{Y}} x_i^2. \tag{8}$$

Since

$$\begin{aligned}
&\sum_{i \in \mathcal{B} \cap \mathcal{Y}} y_i^2 \leq 2 \sum_{i \in \mathcal{B} \cap \mathcal{Y}} (x_i - y_i)^2 + 2 \sum_{i \in \mathcal{B} \cap \mathcal{Y}} x_i^2 \\
\overset{(8)}{\leq} &2 \sum_{i \in \mathcal{B} \cap \mathcal{Y}} (x_i - y_i)^2 + 2 \sum_{i \in \mathcal{A} \setminus \mathcal{Y}} x_i^2,
\end{aligned}$$

we have

$$\begin{aligned}
&\|\mathbf{x}^s - \mathbf{y}\|_2^2 \\
\leq &3 \sum_{i \in \mathcal{A} \setminus \mathcal{Y}} x_i^2 + \sum_{i \in \mathcal{A} \cap \mathcal{Y}} (x_i - y_i)^2 + 2 \sum_{i \in \mathcal{B} \cap \mathcal{Y}} (x_i - y_i)^2 \\
\leq &3 \|\mathbf{x} - \mathbf{y}\|_2^2.
\end{aligned}$$