Supplemental Materials for:
“Solving Random Quadratic Systems of Equations Is Nearly as Easy as Solving Linear Systems”

Yuxin Chen ∗ Emmanuel J. Candès ∗†
May 14, 2015

Abstract
This document presents the proof of the universal stability guarantees for Theorem 2 given in the paper “Solving Random Quadratic Systems of Equations Is Nearly as Easy as Solving Linear Systems”.

1 Universal stability guarantees

In the main text, Theorem 2 has been proved for the situation where the planted solution \( \mathbf{x} \) is fixed independent of the design vectors \( \{ \mathbf{a}_i \} \). This section proves a more universal theory: once the design vectors are selected and fixed, then with high probability TWF succeeds simultaneously for all \( \mathbf{x} \in \mathbb{R}^n \).

Since the main text already delivers universal guarantees for the iterative refinement stage, it only remains to justify universality for truncated spectral initialization. In fact, it suffices to prove that in the absence of noise,

\[
\| \mathbf{Y} - \mathbb{E} [\mathbf{Y}] \| \leq \delta \| \mathbf{x} \|^2, \quad \forall \mathbf{x} \in \mathbb{R}^n
\]  

holds for an arbitrary small constant \( \delta > 0 \), as all remaining steps presented in the main text readily carry over here. To ease presentation, we shall assume \( \| \mathbf{x} \| = 1 \) from now on and denote

\[
\mathbf{Y}_\mathbf{x} := \frac{1}{m} \sum_{i=1}^{m} \mathbf{a}_i \mathbf{a}_i^\top \left( \mathbf{a}_i^\top \mathbf{x} \right)^2 \mathbb{1}_{\{ |\mathbf{a}_i^\top \mathbf{x}| \leq \alpha \}};
\]  

which in expectation gives

\[
\mathbb{E} [\mathbf{Y}_\mathbf{x}] = \beta_1 \mathbf{x} \mathbf{x}^\top + \beta_2 \mathbf{I}.
\]  

Here, \( \beta_1 := \mathbb{E} [\xi^4 \mathbb{1}_{\{ |\xi| \leq \alpha \}}] - \mathbb{E} [\xi^2 \mathbb{1}_{\{ |\xi| \leq \alpha \}}] \) and \( \beta_2 := \mathbb{E} [\xi^2 \mathbb{1}_{\{ |\xi| \leq \alpha \}}] \) where \( \xi \) is a standard normal.

To prove universality, we discretize the unit sphere using an \( \epsilon \)-net \( \mathcal{N}_\epsilon \) of cardinality \( (1 + 2\epsilon) \) so that for any unit vector \( \mathbf{x} \), there exists an \( \mathbf{x}_0 \in \mathcal{N}_\epsilon \) obeying \( \| \mathbf{x} - \mathbf{x}_0 \| \leq \epsilon \). As has been demonstrated in the main text, for any fix \( \mathbf{x}_0 \in \mathbb{R}^n \),

\[
\| \mathbf{Y}_{\mathbf{x}_0} - \mathbb{E} [\mathbf{Y}_{\mathbf{x}_0}] \| \leq \frac{\delta}{2}
\]  

holds with probability \( 1 - \exp (-\Omega (m)) \). Taking the union bound over \( \mathcal{N}_\epsilon \) yields

\[
\| \mathbf{Y}_{\mathbf{x}_0} - \mathbb{E} [\mathbf{Y}_{\mathbf{x}_0}] - \beta_2 \mathbf{I} \| \leq \frac{\delta}{2}, \quad \forall \mathbf{x}_0 \in \mathcal{N}_\epsilon
\]  

with probability \( 1 - \left( 1 + \frac{\delta}{2} \right)^n \exp (-\Omega (m)) \). It then boils down to obtaining uniform control over \( \| \mathbf{Y}_\mathbf{x} - \mathbf{Y}_{\mathbf{x}_0} \| \).

To this end, we decompose \( \mathbf{Y}_\mathbf{x} - \mathbf{Y}_{\mathbf{x}_0} \) in the following manner

\[
\mathbf{Y}_\mathbf{x} - \mathbf{Y}_{\mathbf{x}_0} = \frac{1}{m} \left( \sum_{i \in I_1} + \sum_{i \in I_2} + \sum_{i \in I_3} + \sum_{i \in I_4} \right) \mathbf{a}_i \mathbf{a}_i^\top \cdot \left( (\mathbf{a}_i^\top \mathbf{x})^2 \mathbb{1}_{\{ |\mathbf{a}_i^\top \mathbf{x}| \leq \alpha \}} - (\mathbf{a}_i^\top \mathbf{x}_0)^2 \mathbb{1}_{\{ |\mathbf{a}_i^\top \mathbf{x}_0| \leq \alpha \}} \right);
\]  

∗Department of Statistics, Stanford University, Stanford, CA 94305, U.S.A.
†Department of Mathematics, Stanford University, Stanford, CA 94305, U.S.A.
Consequently, one can write
\[ \|a_i^\top x\| \leq \|a_i^\top x_0\| + \zeta \leq \alpha_y, \]
and hence both indicator variables \( 1\{a_i^\top x \leq \alpha_y\} \) and \( 1\{a_i^\top x_0 \leq \alpha_y\} \) are active. In addition, the constraint that specifies \( I_1 \) gives
\[ \|a_i^\top x - (a_i^\top x_0)^2\| = \|a_i^\top (x - x_0)\| \cdot |a_i^\top x + a_i^\top x_0| \leq \zeta \cdot 2\alpha_y. \]
Consequently, one can write
\[
\left\| \frac{1}{m} \sum_{i \in I_1} a_i a_i^\top \right\| \left\{ (a_i^\top x)^2 1\{a_i^\top x \leq \alpha_y\} - (a_i^\top x_0)^2 1\{a_i^\top x_0 \leq \alpha_y\} \right\} = \frac{1}{m} \sum_{i \in I_1} a_i a_i^\top \left( (a_i^\top x)^2 - (a_i^\top x_0)^2 \right) \\
\leq 2\zeta \|a_i^\top x\| \leq 2\zeta |a_i^\top x_0 - \zeta|, \\
\leq 2\zeta |a_i^\top x_0 - \zeta| \leq 2\zeta \alpha_y, \quad \text{(6)}
\]
provided that \( \|\frac{1}{m} A^\top A\| \leq 2 \) (which occurs with probability \( 1 - \exp(-\Omega(m)) \)).

**Case 2.** When \( i \in I_2 \), both indicator variables are zero since
\[ |a_i^\top x| \geq |a_i^\top x_0| - \zeta > \alpha_y, \]
leading to
\[
\frac{1}{m} \sum_{i \in I_2} a_i a_i^\top \left\{ (a_i^\top x)^2 1\{a_i^\top x \leq \alpha_y\} - (a_i^\top x_0)^2 1\{a_i^\top x_0 \leq \alpha_y\} \right\} = 0.
\]

**Case 3 and Case 4.** The index sets \( I_3 \) and \( I_4 \) represent the region where \( 1\{a_i^\top x \leq \alpha_y\} \) and \( 1\{a_i^\top x_0 \leq \alpha_y\} \) might disagree. For these two cases, one can only bound
\[
\left\| \frac{1}{m} \sum_{i \in I_3 \cup I_4} a_i a_i^\top \left\{ (a_i^\top x)^2 1\{a_i^\top x \leq \alpha_y\} - (a_i^\top x_0)^2 1\{a_i^\top x_0 \leq \alpha_y\} \right\} \right\| \leq 2\alpha_y \left\| \frac{1}{m} \sum_{i \in I_3 \cup I_4} a_i a_i^\top \right\| \quad \text{(7)}
\]
using the truncation rule \( |a_i^\top x| \leq \alpha_y \) and \( |a_i^\top x_0| \leq \alpha_y \). Fortunately, \( I_3 \) and \( I_4 \) correspond to a collection of rare events. In fact, for any small constant \( \zeta > 0 \), the standard concentration inequality together with the union bound give
\[
\frac{1}{m} \sum_{i=1}^m 1\{a_i^\top x_0 \in [\alpha_y \pm \zeta]\} \leq (1 + \epsilon) \mathbb{E} \left[ 1\{a_i^\top x_0 \in [\alpha_y \pm \zeta]\} \right], \quad \forall x_0 \in \mathcal{N},
\]
with probability at least \( 1 - (1 + \frac{2}{\zeta})^n \exp(-\Omega(m)) \). When \( \zeta \) is sufficiently small, this suggests
\[
\frac{1}{m} \sum_{i=1}^m 1\{a_i^\top x_0 \in [\alpha_y \pm \zeta]\} \leq \frac{\vartheta}{2}, \quad \forall x_0 \in \mathcal{N},
\]
for some small constant \( \vartheta > 0 \). Additionally, it follows from Lemma 6 in the main text that with high probability,
\[
\frac{1}{m} \sum_{i=1}^m 1\{a_i^\top (x - x_0) \geq \zeta\} \leq \frac{\vartheta}{2}, \quad \forall x \in \mathbb{R}^n,
\]
provided that the separation $\epsilon$ of the $\epsilon$-net is small enough. In summary, the cardinality of the index sets $I_3$ and $I_4$ combined together cannot exceed $\vartheta m$.

It then comes down to controlling the spectral norm of $A_S$ for all index sets $S$ obeying $|S| \leq \vartheta m$, where $A_S$ denotes the submatrix of $A$ comprising $\{a_i\}_{i \in S}$. Standard random matrix theory (e.g., [1, Corollary 5.35]) suggests that for any given $S$ obeying $|S| \leq \vartheta m$,

$$\|A_S\| \leq \sqrt{\vartheta m} + \sqrt{n} + \tau$$

holds with probability exceeding $1 - 2\exp(-\frac{\tau^2}{2})$. Note that the total number of index sets $S$ with $|S| \leq \vartheta m$ is bounded above by [2, Example 11.1.3]

$$(m_{|S|}) \leq e^{m_{\mathcal{H}(\vartheta)}},$$

with $\mathcal{H}(\vartheta) := -\vartheta \log \vartheta - (1 - \vartheta) \log(1 - \vartheta)$ denoting the binary entropy function. Setting $\tau = 2\sqrt{\mathcal{H}(\vartheta)m}$ we get

$$\frac{1}{\sqrt{m}} \|A_S\| \leq \sqrt{\vartheta} + \sqrt{\frac{n}{m}} + 2\sqrt{\mathcal{H}(\vartheta)}, \quad \forall S : |S| \leq \vartheta m.$$ 

with probability at least $1 - \exp(-\Omega(m))$. Thus, for any constant $\delta > 0$, one has

$$\frac{1}{\sqrt{m}} \|A_S\| \leq \tilde{\delta}, \quad \forall S : |S| \leq \vartheta m$$

as long as $m/n$ is sufficiently large and $\vartheta$ is sufficiently small. This combined with (7) yields

$$\left\| \frac{1}{m} \sum_{i \in I_3 \cup I_4} a_i a_i^\top \left\{ (a_i^\top x)^2 1_{\{|a_i^\top x| \leq \alpha_y \}} - (a_i^\top x_0)^2 1_{\{|a_i^\top x_0| \leq \alpha_y \}} \right\} \right\| \leq \frac{2\alpha_y^2}{m} \|A_{I_3 \cup I_4}\|^2 \leq 2\alpha_y^2 \tilde{\delta}^2.$$  

(8)

To finish up, we put the above cases together to deduce

$$\left\| \frac{1}{m} \sum_{i=1}^m a_i a_i^\top \left\{ (a_i^\top x)^2 1_{\{|a_i^\top x| \leq \alpha_y \}} - (a_i^\top x_0)^2 1_{\{|a_i^\top x_0| \leq \alpha_y \}} \right\} \right\| \leq 4\zeta_a + 2\alpha_y^2 \tilde{\delta}^2,$$

(9)

and hence

$$\|Y_x - \beta_1 xx^\top - \beta_2 I\| \leq \|Y_x - Y_{x_0}\| + \|Y_{x_0} - \beta_1 x_0 x_0^\top - \beta_2 I\| + \beta_1 \|x_0 x_0^\top - xx^\top\|$$

$$\leq 4\zeta_a + 2\alpha_y^2 \tilde{\delta}^2 + \frac{\delta}{2} + 2.5\beta_1 \epsilon,$$

(10)

(11)

where the last inequality makes use of Lemma 2 of the main text, i.e., $\|x_0 x_0^\top - xx^\top\| \leq 2.5\|x - x_0\|\|x\| \leq 2.5\epsilon$. Since $\tilde{\delta}, \zeta, \epsilon$ can all be arbitrarily small, this establishes [1] and in turn the universal guarantees.

References
