Stability and robustness of ℓ_1-minimizations with Weibull matrices and redundant dictionaries[*†]

Simon Foucart[‡], Drexel University

Abstract

We investigate the recovery of almost s-sparse vectors $\mathbf{x} \in \mathbb{C}^N$ from undersampled and inaccurate data $\mathbf{y} = A\mathbf{x} + \mathbf{e} \in \mathbb{C}^m$ by means of minimizing $\|\mathbf{z}\|_1$ subject to the equality constraints $A\mathbf{z} = \mathbf{y}$. If $m \asymp s\ln(N/s)$ and if Gaussian random matrices $A \in \mathbb{R}^{m\times N}$ are used, this equality-constrained ℓ_1-minimization is known to be stable with respect to sparsity defects and robust with respect to measurement errors. If $m \asymp s\ln(N/s)$ and if Weibull random matrices are used, we prove here that the equality-constrained ℓ_1-minimization remains stable and robust. The arguments are based on two key ingredients, namely the robust null space property and the quotient property. The robust null space property relies on a variant of the classical restricted isometry property where the inner norm is replaced by the ℓ_1-norm and the outer norm is replaced by a norm comparable to the ℓ_2-norm. For the ℓ_1-minimization subject to inequality constraints, this yields stability and robustness results that are also valid when considering sparsity relative to a redundant dictionary. As for the quotient property, it relies on lower estimates for the tail probability of sums of independent Weibull random variables.

1 Background and New Contributions

This article deals with the approximate recovery of signals using ℓ_1-minimizations. In classical works, the signals $\mathbf{f}$ are assumed to be compressible in an orthonormal basis, and they are identified with their vectors $\mathbf{x} \in \mathbb{C}^N$ of coefficients on this orthonormal basis. The signals are then acquired via the measurement vectors $\mathbf{y} = A\mathbf{x} + \mathbf{e} \in \mathbb{C}^m$, $m \ll N$, where $A \in \mathbb{C}^{m\times N}$ is a measurement matrix (the same for all signals) and $\mathbf{e} \in \mathbb{C}^m$ are error vectors. If the measurement matrix A satisfies a restricted isometry condition of order s and if an estimate $\|\mathbf{e}\|_2 \le \eta$ is available, then the reconstruction map given by the inequality-constrained ℓ_1-minimization

$$\Delta_{1,\eta}(\mathbf{y}) := \underset{\mathbf{z}\in\mathbb{C}^N}{\operatorname{argmin}} \|\mathbf{z}\|_1 \quad \text{subject to } \|A\mathbf{z} - \mathbf{y}\|_2 \le \eta$$

yields bounds on the reconstruction error of the type

$$\|\mathbf{x} - \Delta_{1,\eta}(A\mathbf{x} + \mathbf{e})\|_1 \le C\,\sigma_s(\mathbf{x})_1 + D\sqrt{s}\,\eta, \tag{1}$$

$$\|\mathbf{x} - \Delta_{1,\eta}(A\mathbf{x} + \mathbf{e})\|_2 \le \frac{C}{\sqrt{s}}\,\sigma_s(\mathbf{x})_1 + D\,\eta, \tag{2}$$

[*]Key words: compressive sensing, ℓ_1-minimization, basis pursuit, robust null space property, quotient property, Weibull random variables, frames and redundant dictionaries

[†]AMS classification: 42C15, 65K10, 94A12, 94A20

[‡]The author is supported by the NSF grant DMS-1120622

see [3] for (2), [4] for (1) with $\eta = 0$, and [7] for a general statement valid for $p \in [1, 2]$. Here,

$$\sigma_s(\mathbf{x})_1 := \min\{\|\mathbf{x} - \mathbf{z}\|_1, \mathrm{card}(\mathrm{supp}(\mathbf{z})) \leq s\}$$

denotes the error of best ℓ_1-approximation of $\mathbf{x}$ by s-sparse vectors (i.e., vectors with at most s nonzero entries). In particular, for exactly s-sparse vectors (i.e., $\sigma_s(\mathbf{x})_1 = 0$) measured with perfect accuracy (i.e., $\eta = 0$), the reconstruction is completely faithful (that is, $\Delta_{1,0}(A\mathbf{x}) = \mathbf{x}$). This framework has recently been generalized to signals that are compressible in a tight frame, which is represented by a matrix $B \in \mathbb{C}^{n \times N}$ satisfying $BB^* = \mathrm{Id}$. Namely, signals $\mathbf{f} \in \mathbb{C}^n$ are obtained from their vectors $\mathbf{x} \in \mathbb{C}^N$ of coefficients as

$$\mathbf{f} = B\mathbf{x} \in \mathbb{C}^n, \qquad \text{where} \qquad \mathbf{x} := B^*\mathbf{f} \in \mathbb{C}^N.$$

The signals are acquired via the measurement vectors $\mathbf{y} = A\mathbf{f} + \mathbf{e} = AB\mathbf{x} + \mathbf{e} \in \mathbb{C}^m$ for some measurement matrix $A \in \mathbb{C}^{m \times n}$ with $m < n$ and some error vectors $\mathbf{e} \in \mathbb{C}^m$. The bound (2) was extended in [2] to ensure that, if $m \geq c\, s \ln(eN/s)$, then with high probability on the draw of a subgaussian random matrix $A \in \mathbb{C}^{m \times n}$, the bound

$$\|B^*(\mathbf{f} - \Lambda_{1,\eta}(A\mathbf{f} + \mathbf{e}))\|_2 \leq \frac{C}{\sqrt{s}}\, \sigma_s(B^*\mathbf{f})_1 + D\,\eta \tag{3}$$

holds for all $\mathbf{f} \in \mathbb{C}^n$ and $\mathbf{e} \in \mathbb{C}^m$ with $\|\mathbf{e}\|_2 \leq \eta$, where the reconstructed vector $\Lambda_{1,\eta}(A\mathbf{f} + \mathbf{e})$ is a minimizer of $\|B^*\mathbf{g}\|_1$ subject to $\|A\mathbf{g} - \mathbf{y}\|_2 \leq \eta$. The bound (3) was established in the real setting and was in fact stated for $\|\mathbf{f} - \Lambda_{1,\eta}(A\mathbf{f} + \mathbf{e})\|_2$, but the formulations are equivalent for tight frames, since

$$\|B^*\mathbf{g}\|_2^2 = \langle B^*\mathbf{g}, B^*\mathbf{g}\rangle = \langle \mathbf{g}, BB^*\mathbf{g}\rangle = \langle \mathbf{g}, \mathbf{g}\rangle = \|\mathbf{g}\|_2^2 \qquad \text{for all } \mathbf{g} \in \mathbb{C}^n.$$

In our considerations, we may quantify the measurement error via $\|\mathbf{e}\| \leq \eta$ for some other norm than the ℓ_2-norm, and we may replace B^* by an unspecified $B' \in \mathbb{C}^{N \times n}$, so that the reconstruction map is given by the inequality-constrained ℓ_1-minimization

$$\Lambda_{1,\eta}(\mathbf{y}) := \underset{\mathbf{g} \in \mathbb{C}^n}{\mathrm{argmin}}\ \|B'\mathbf{g}\|_1 \quad \text{subject to } \|A\mathbf{g} - \mathbf{y}\| \leq \eta. \tag{4}$$

In both basis and frame cases, the arguments are based on the restricted isometry property of order s, which is typically fulfilled with high probability on the draw of an isotropic subgaussian random matrix $A \in \mathbb{C}^{m \times N}$ provided $m \geq c\, s \ln(eN/s)$. In addition, for Gaussian random matrices in the basis case, it is even known [15] that a prior estimation η of the measurement error is not necessary to obtain bounds comparable to (2). Indeed, with high probability on the draw of a Gaussian matrix $A \in \mathbb{C}^{m \times N}$, the reconstruction map given by the equality-constrained ℓ_1-minimization[1]

$$\Delta_1(\mathbf{y}) := \underset{\mathbf{z} \in \mathbb{C}^N}{\mathrm{argmin}}\ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}$$

[1] When considering this minimization problem, it is implicitly assumed that it is feasible for all $\mathbf{y} \in \mathbb{C}^m$, i.e., that A is surjective.

yields a bound on the reconstruction error of the type

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_2 \le \frac{C}{\sqrt{s}}\,\sigma_s(\mathbf{x})_1 + D\,\|\mathbf{e}\|_2 \tag{5}$$

holding for all $\mathbf{x} \in \mathbb{C}^N$ and all $\mathbf{e} \in \mathbb{C}^m$, as long as $m \ge c\,s\ln(eN/s)$ (the additional condition $N \ge m\ln^2(m)$, required in [15], is in fact not necessary). The case of Bernoulli matrices was studied in [5], where a variation of (5) was established by altering the ℓ_2-norm on the right-hand side. The results of [15, 5] were further complemented in [16] for matrices satisfying the restricted isometry property. But for other matrices with $m \asymp s\ln(eN/s)$ (for instance, for matrices populated by independent Laplace random variables), the restricted isometry property is not fulfilled [1], yet faithful reconstruction of exactly sparse vectors measured with perfect accuracy via ℓ_1-minimization can be achieved, see [8].

The main objective of this article is to prove stability and robustness estimates of types (3) and (5) for matrices without the restricted isometry property in the standard form. The article reveals in particular that Gaussian matrices are not the only ones providing the estimate (5) for equality-constrained ℓ_1-minimization. Weibull random matrices are used for this purpose. The entries $a_{i,j}$ of such matrices $A \in \mathbb{R}^{m\times N}$ are independent symmetric Weibull random variables with exponent $r \ge 1$, i.e., their tail probabilities satisfy

$$\Pr(|a_{i,j}| \ge t) = \exp\left(-\left(\frac{\sqrt{\Gamma(1+2/r)}\,t}{\sigma_r}\right)^r\right), \quad t \ge 0.$$

These symmetric Weibull random variables have mean zero and variance σ_r^2. The value $r = 1$, for instance, corresponds to Laplace random variables, and the value $r = \infty$ to Rademacher random variables. Our definition of Weibull matrices requires that the squared ℓ_2-norm of each column has mean equal to one, which we ensure by imposing

$$\sigma_r = \frac{1}{\sqrt{m}}.$$

We point out that the first absolute moment $\mu_r := \mathbb{E}|a_{i,j}|$ of the random variables $a_{i,j}$ is then

$$\mu_r = \frac{\Gamma(1+1/r)}{\sqrt{\Gamma(1+2/r)}}\frac{1}{\sqrt{m}}, \qquad \text{so that} \qquad \mu_r \ge \frac{1}{\sqrt{2}\sqrt{m}}.$$

The article centers on the following two theorems. The first theorem incorporate estimates similar to (3) as the particular case $p = 2$. Here, the matrix $B' \in \mathbb{C}^{N\times n}$ in the definition (4) of the reconstruction map $\Lambda_{1,\eta}$ equals the Moore–Penrose pseudo-inverse

$$B^\dagger := B^*(BB^*)^{-1}.$$

Theorem 1. Let $B \in \mathbb{C}^{n\times N}$, $n \le N$, be a matrix with smallest and largest singular values $\sigma_{\min}$ and $\sigma_{\max}$. If $A \in \mathbb{R}^{m\times n}$ is a Weibull random matrix with exponent $1 \le r \le 2$, then with probability at least $1 - 2\exp(-c_1 m)$, the bounds

$$\|B^\dagger(\mathbf{f} - \Lambda_{1,\eta}(A\mathbf{f} + \mathbf{e}))\|_p \le \frac{C}{s^{1-1/p}}\,\sigma_s(B^\dagger\mathbf{f})_1 + D\,s^{1/p-1/2}\,\eta, \qquad 1 \le p \le 2, \tag{6}$$

hold for all $\mathbf{f} \in \mathbb{C}^n$ and $\mathbf{e} \in \mathbb{C}^m$ with $\|\mathbf{e}\|_2 \le \eta$, as long as $m \ge c_2\,s\ln(eN/s)$. The constants $c_1, C > 0$ are universal, while the constants $c_2, D > 0$ depend on the ratio $\sigma_{\max}/\sigma_{\min}$.

Theorem 1 extends the results of [2] to matrices that are not subgaussian. It also applies to subgaussian matrices, and the statement for frames that are not necessarily tight, for p that does not necessarily equal 2, and for the complex setting rather than the real setting are novel in this context, too. The canonical basis case $B = \mathrm{Id}$ of Theorem 1 is needed to establish the second theorem, which is concerned with the equality-constrained ℓ_1-minimization.

Theorem 2. If $A \in \mathbb{R}^{m \times N}$ is a Weibull random matrix with exponent $1 \leq r \leq 2$, then with probability at least $1 - 3\exp(-c_1 m)$, the bounds

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 + D\, s_*^{1/p-1/2}\, \|\mathbf{e}\|_2, \qquad 1 \leq p \leq 2, \tag{7}$$

are valid for all $\mathbf{x} \in \mathbb{C}^N$ and all $\mathbf{e} \in \mathbb{C}^m$, as long as $N \geq c_2 m$ and $s \leq c_3\, s_*$, $s_* := m/\ln(eN/m)$. The constants $c_1, c_2, c_3, C, D > 0$ are absolute constants.

It is natural to wonder if Theorem 2 extends from the basis case to the frame case, but this question provides additional challenges, and it is not addressed here. The article is divided in three remaining sections. In Section 2, two properties that must be fulfilled for the estimates of Theorems 1 and 2 to hold are isolated: the robust null space property and the quotient property. In the course of the article, these necessary properties are shown to be sufficient, too. Sections 3 and 4 focus on the inequality-constrained and on the equality-constrained ℓ_1-minimizations, respectively. They are organized on the same model: a first subsection based on deterministic arguments proves the sufficiency of one of the two properties, it is then followed by a subsection containing auxiliary results, and finally probabilistic arguments are used to show that the property under consideration holds for random matrices. Although the attention is put on Weibull random matrices, a subsection on subgaussian random matrices is also included in Section 3.

2 Two Necessary Properties

The two ingredients at the core of Theorems 1 and 2 are the robust null space property and the quotient property. Our strategy consists in proving that the robust null space property guarantees the stability and robustness of the inequality-constrained ℓ_1-minimization, that the combination of the robust null space property and the quotient property guarantees the stability and robustness of the equality-constrained ℓ_1-minimization, and that the two properties are satisfied with high probability for Weibull random matrices. As an intuitive way to understand the two properties, it is worth pointing out that they must be fulfilled for the prospective estimates to hold.

Firstly, considering Theorem 1 where we deliberately replace in (6) the pseudo-inverse $B^\dagger$ by an unspecified right-inverse B' and the norm $\|\mathbf{e}\|_2$ by an unspecified norm $\|\mathbf{e}\|$, the prospective bound reads

$$\|B'(\mathbf{f} - \Lambda_{1,\eta}(A\mathbf{f} + \mathbf{e}))\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(B'\mathbf{f})_1 + D\, s^{1/p-1/2}\, \eta \tag{8}$$

for all $\mathbf{f} \in \mathbb{C}^n$ and all $\mathbf{e} \in \mathbb{C}^m$ with $\|\mathbf{e}\| \leq \eta$. Given $\mathbf{v} \in \mathbb{C}^n$, if we set $\mathbf{f} = \mathbf{v}$, $\mathbf{e} = -A\mathbf{v}$, and $\eta = \|A\mathbf{v}\|$, so that $B'\Lambda_{1,\eta}(A\mathbf{f} + \mathbf{e}) = 0$, then (8) yields

$$\|B'\mathbf{v}\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(B'\mathbf{v})_1 + D\, s^{1/p-1/2}\, \|A\mathbf{v}\|.$$

For any index set S of size at most s, the inequalities $\|(B'\mathbf{v})_S\|_p \leq \|B'\mathbf{v}\|_p$ and $\sigma_s(B'\mathbf{v})_1 \leq \|(B'\mathbf{v})_{\overline{S}}\|_1$ now give

$$\|(B'\mathbf{v})_S\|_p \leq \frac{C}{s^{1-1/p}}\, \|(B'\mathbf{v})_{\overline{S}}\|_1 + D\, s^{1/p-1/2}\, \|A\mathbf{v}\|.$$

This is the ℓ_p-robust null space property of order s relative to the norm $s^{1/p-1/2}\, \|\cdot\|$. It could be seen that, in the canonical basis case where $B' = \mathrm{Id}$, this property also follows from the prospective estimate (7) by taking $\mathbf{x} = \mathbf{v}$ and $\mathbf{e} = -A\mathbf{v}$.

Definition 3. Given $p \geq 1$, the matrices $A \in \mathbb{C}^{m\times n}$ and $B' \in \mathbb{C}^{N\times n}$ are said to satisfy the ℓ_p-robust null space property of order s with constant $\rho > 0$ and $\tau > 0$ relative to a norm $\|\cdot\|$ on $\mathbb{C}^m$ if

$$\text{(9)}\quad \|(B'\mathbf{v})_S\|_p \leq \frac{\rho}{s^{1-1/p}}\, \|(B'\mathbf{v})_{\overline{S}}\|_1 + \tau\, \|A\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{C}^n \text{ and all } S \subseteq [N] \text{ with } \mathrm{card}(S) \leq s.$$

The requirement $\rho < 1$ is implicitly added in informal mentions of the robust null space property. This requirement is best understood by specifying the ℓ_1-robust null space property to $\mathbf{v} \in \ker A$ to obtain the null space property

$$\text{(10)}\qquad \|(B'\mathbf{v})_S\|_1 < \|(B'\mathbf{v})_{\overline{S}}\|_1 \qquad \text{for all } \mathbf{v} \in \ker A \setminus \{0\} \text{ and all } S \subseteq [N] \text{ with } \mathrm{card}(S) \leq s.$$

This condition easily implies that every $\mathbf{f} \in \mathbb{C}^n$ such that $B'\mathbf{f}$ is s-sparse uniquely minimizes $\|B'\mathbf{g}\|_1$ subject to $A\mathbf{g} = A\mathbf{f}$. Indeed, if $\mathbf{g} \in \mathbb{C}^n$, $\mathbf{g} \neq \mathbf{f}$, is such that $A\mathbf{g} = A\mathbf{f}$, then (10) applied to $\mathbf{v} := \mathbf{f} - \mathbf{g} \in \ker A \setminus \{0\}$ and $S = \mathrm{supp}(B'\mathbf{f})$ yields

$$\begin{aligned}\|B'\mathbf{g}\|_1 &= \|(B'\mathbf{g})_S\|_1 + \|(B'\mathbf{g})_{\overline{S}}\|_1 = \|B'\mathbf{f} - (B'\mathbf{v})_S\|_1 + \|(B'\mathbf{v})_{\overline{S}}\|_1 \geq \|B'\mathbf{f}\|_1 - \|(B'\mathbf{v})_S\|_1 + \|(B'\mathbf{v})_{\overline{S}}\|_1\\ &> \|B'\mathbf{f}\|_1.\end{aligned}$$

In the basis case where $B' = B^{-1}$, the converse is well-known. Precisely, the unique recovery via ℓ_1-minimization of every $\mathbf{f} \in \mathbb{C}^n$ such that $B'\mathbf{f}$ is s-sparse implies the null space property. Indeed, for $\mathbf{v} \in \ker A$ and $S \subseteq [N]$ with $\mathrm{card}(S) = s$, the equality $0 = A\mathbf{v} = ABB'\mathbf{v}$ yields $A\mathbf{f} = A\mathbf{g}$ with $\mathbf{f} := B(B'\mathbf{v})_S$ and $\mathbf{g} := -B(B'\mathbf{v})_{\overline{S}}$, but since the vector $B'\mathbf{f} = (B'\mathbf{v})_S$ is s-sparse, it must indeed have a smaller ℓ_1-norm than $B'\mathbf{g} = -(B'\mathbf{v})_{\overline{S}}$.

Secondly, considering Theorem 2 where we deliberately replace in (7) the norm $\|\mathbf{e}\|_2$ by an unspecified norm $\|\mathbf{e}\|$, the prospective bound reads

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 + D\, s_*^{1/p-1/2}\, \|\mathbf{e}\|.$$

If we first set $\mathbf{x} = 0$, so that $\Delta_1(A\mathbf{f} + \mathbf{e}) = \Delta_1(\mathbf{e})$, this bound taken for $p = 1$ yields

$$\|\Delta_1(\mathbf{e})\|_1 \leq D\,\sqrt{s_*}\,\|\mathbf{e}\|,$$

which is equivalent to the existence of $\mathbf{v} \in \mathbb{C}^N$ with $A\mathbf{v} = \mathbf{e}$ and $\|\mathbf{v}\|_1 \leq D\sqrt{s_*}\|\mathbf{e}\|$, or to an equality between the norm of $\mathbf{e}$ and the norm of its set of preimages in the quotient space of ℓ_1^N by $\ker A$, namely

$$\|[\mathbf{e}]\|_A := \inf\{\|\mathbf{v}\|_1, A\mathbf{v} = \mathbf{e}\} \leq D\,\sqrt{s_*}\,\|\mathbf{e}\|.$$

This explains the terminology ℓ_1-quotient property. As a generalization to $p \geq 1$, we introduce the ℓ_p-quotient property as follows.[2]

Definition 4. Given $p \geq 1$, the matrix $A \in \mathbb{C}^{m\times N}$ is said to satisfy the ℓ_p-quotient property with constant d relative to a norm $\|\cdot\|$ on $\mathbb{C}^m$ if, with $s_* := m/\ln(eN/m)$,

$$\text{for all } \mathbf{e} \in \mathbb{C}^m, \text{ there exists } \mathbf{v} \in \mathbb{C}^N \text{ with } A\mathbf{v} = \mathbf{e} \text{ and } \|\mathbf{v}\|_p \leq d\,s_*^{1/p-1/2}\,\|\mathbf{e}\|.$$

3 Inequality-Constrained ℓ_1-Minimization

This section is devoted to the proof of Theorem 1, whose special case $B = \mathrm{Id}$ will be needed in Section 4. The deterministic argument showing that the estimate (6) is a consequence of the robust null space property are presented in Subsection 3.1. Some preliminary lemmas involved in the probabilistic justification that random matrices satisfy the robust null space property are collected in Subsection 3.2. The case of subgaussian matrices is then revisited in Subsection 3.3. Finally, Subsection 3.4 deals with Weibull matrices, and more generally matrices whose entries are independent pregaussian (more often called ψ_1) random variables.

3.1 Sufficiency of the Robust Null Space Property

The result below outlines the relevance of the ℓ_p-robust null space property when generalizing estimate (3) to $p \neq 2$. It is stated for an unspecified $B' \in \mathbb{C}^{N\times n}$ rather than the pseudo-inverse B^* of a tight frame $B \in \mathbb{C}^{n\times N}$.

Theorem 5. Given $q \geq 1$, suppose that the matrices $A \in \mathbb{C}^{m\times n}$ and $B' \in \mathbb{C}^{N\times n}$ satisfy the ℓ_q-robust null space property of order s with constants $0 < \rho < 1$ and $\tau > 0$ relative to a norm $\|\cdot\|$ on $\mathbb{C}^m$. Then, for any $1 \leq p \leq q$, the bounds

$$\|B'(\mathbf{f} - \Lambda_{1,\eta}(A\mathbf{f} + \mathbf{e}))\|_p \leq \frac{C}{s^{1-1/p}}\,\sigma_s(B'\mathbf{f})_1 + D\,s^{1/p-1/q}\,\eta,$$

hold for all $\mathbf{f} \in \mathbb{C}^n$ and $\mathbf{e} \in \mathbb{C}^m$ with $\|\mathbf{e}\| \leq \eta$. The constant $C, D > 0$ depend only on ρ and τ.

---

[2] Note in particular that a matrix with the quotient property is automatically sujective.

Proof. We shall prove the stronger statement that, for any $\mathbf{f}, \mathbf{g} \in \mathbb{C}^n$,

$$\|B'(\mathbf{f}-\mathbf{g})\|_p \le \frac{C}{s^{1-1/p}} \left(\|B'\mathbf{g}\|_1 - \|B'\mathbf{f}\|_1 + 2\sigma_s(B'\mathbf{f})_1 \right) + D\, s^{1/p-1/q}\, \|A(\mathbf{f}-\mathbf{g})\|. \tag{11}$$

Then, choosing $\mathbf{g} := \Lambda_{1,\eta}(A\mathbf{f}+\mathbf{e})$, the announced result follows (with different constants) from the inequality $\|B'\mathbf{g}\|_1 \le \|B'\mathbf{f}\|_1$ and the inequality $\|A(\mathbf{f}-\mathbf{g})\| \le \|A\mathbf{f}-\mathbf{y}\| + \|A\mathbf{g}-\mathbf{y}\| \le 2\eta$, where $\mathbf{y} = A\mathbf{f}+\mathbf{e}$. We notice first that, for $1 \le p \le q$, the ℓ_q-robust null space property of order s relative to $\|\cdot\|$ implies the ℓ_p-robust null space property of order s relative to $s^{1/p-1/q}\,\|\cdot\|$, by virtue of the inequality $\|(B'\mathbf{v})_S\|_p \le s^{1/p-1/q}\,\|(B'\mathbf{v})_S\|_q$. Let us now isolate the case $p=1$. Here, the ℓ_1-robust null space property applied to the vector $\mathbf{v} := \mathbf{g}-\mathbf{f} \in \mathbb{C}^n$ and to an index set $S \subseteq [N]$ of s largest absolute entries of $B'\mathbf{v}$ reads

$$\|(B'\mathbf{v})_S\|_1 \le \rho\, \|(B'\mathbf{v})_{\overline{S}}\|_1 + \tau\, s^{1-1/q}\, \|A\mathbf{v}\|. \tag{12}$$

We write

$$\begin{aligned}
\|B'\mathbf{f}\|_1 &= \|(B'\mathbf{f})_S\|_1 + \|(B'\mathbf{f})_{\overline{S}}\|_1 \le \|(B'\mathbf{v})_S\|_1 + \|(B'\mathbf{g})_S\|_1 + \|(B'\mathbf{f})_{\overline{S}}\|_1,\\
\|(B'\mathbf{v})_{\overline{S}}\|_1 &\le \|(B'\mathbf{g})_{\overline{S}}\|_1 + \|(B'\mathbf{f})_{\overline{S}}\|_1.
\end{aligned}$$

After summation and rearrangement, we obtain

$$\|(B'\mathbf{v})_{\overline{S}}\|_1 \le \|(B'\mathbf{v})_S\|_1 + \|B'\mathbf{g}\|_1 - \|B'\mathbf{f}\|_1 + 2\|(B'\mathbf{f})_{\overline{S}}\|_1. \tag{13}$$

Combining (13) and (12) gives, after rearrangement,

$$\|(B'\mathbf{v})_{\overline{S}}\|_1 \le \frac{1}{1-\rho}\Bigg(\|B'\mathbf{g}\|_1 - \|B'\mathbf{f}\|_1 + 2\|(B'\mathbf{f})_{\overline{S}}\|_1 + \tau\, s^{1-1/q}\, \|A\mathbf{v}\| \Bigg). \tag{14}$$

Using (12) once again, together with (14), we deduce

$$\begin{aligned}
\|B'\mathbf{v}\|_1 &= \|(B'\mathbf{v})_{\overline{S}}\|_1 + \|(B'\mathbf{v})_S\|_1 \le (1+\rho)\|(B'\mathbf{v})_{\overline{S}}\|_1 + \tau\, s^{1-1/q}\, \|A\mathbf{v}\|\\
&\le \frac{1+\rho}{1-\rho}\Bigg(\|B'\mathbf{g}\|_1 - \|B'\mathbf{f}\|_1 + 2\|(B'\mathbf{f})_{\overline{S}}\|_1 \Bigg) + \frac{2\tau}{1-\rho}\, s^{1-1/q}\, \|A\mathbf{v}\|.
\end{aligned} \tag{15}$$

With $C := (1+\rho)/(1-\rho)$ and $D := 2\tau/(1-\rho)$, we recognize the stated inequality (11) for $p=1$. Interestingly, it could be conversely verified that (11) for $p=1$ implies the ℓ_1-robust null space property. Now, for the general case $1 \le p \le q$, we use the classical inequality

$$\sigma_s(B'\mathbf{v})_p \le \frac{1}{s^{1-1/p}} \|B'\mathbf{v}\|_1,$$

as well as the ℓ_p-robust null space property. Thus, we obtain

$$\begin{aligned}
\|B'\mathbf{v}\|_p &\le \|(B'\mathbf{v})_{\overline{S}}\|_p + \|(B'\mathbf{v})_S\|_p \le \frac{1}{s^{1-1/p}}\|B'\mathbf{v}\|_1 + \frac{\rho}{s^{1-1/p}}\, \|(B'\mathbf{v})_{\overline{S}}\|_1 + \tau\, s^{1/p-1/q}\, \|A\mathbf{v}\|\\
&\le \frac{1+\rho}{s^{1-1/p}}\, \|B'\mathbf{v}\|_1 + \tau\, s^{1/p-1/q}\, \|A\mathbf{v}\|.
\end{aligned}$$

It remains to substitute (15) into the latter to deduce the desired inequality (11). $\square$

Remark. If we impose $\eta = 0$, thus considering the reconstruction map given by the equality-constrained ℓ_1-minimization

$$\Lambda_1(\mathbf{y}) := \operatorname*{argmin}_{\mathbf{g}\in\mathbb{C}^n} \|B'\mathbf{g}\|_1 \quad \text{subject to } A\mathbf{g} = \mathbf{y},$$

we see that if A and B' satisfy the ℓ_p-robust null space property of order s, then the triple (A, B', Λ_1) is mixed (ℓ_p, ℓ_1)-instance optimal of order s with constant $C > 0$ in the sense of [4], i.e.,

$$\|B'(\mathbf{f} - \Lambda_1(A\mathbf{f}))\|_p \le \frac{C}{s^{1-1/p}} \sigma_s(B'\mathbf{f})_1 \qquad \text{for all } \mathbf{f} \in \mathbb{C}^n. \tag{16}$$

3.2 Auxiliary Lemmas

We now isolate two simple results used in the next two subsections. The first result is an observation concerning a modified form of the robust null space property. This form emerges more naturally than the original form of the robust null space property in our upcoming proofs.

Lemma 6. Given $p \ge 1$, the matrices $A \in \mathbb{C}^{m\times n}$ and $B' \in \mathbb{C}^{N\times n}$ satisfy the ℓ_p-robust null space property of order s for some constants $0 < \rho < 1$ and $\tau > 0$ relative to a norm $\|\cdot\|$ as soon as there exist constants $0 < \rho' < 1/2$ and $\tau' > 0$ such that

$$\|(B'\mathbf{v})_S\|_p \le \frac{\rho'}{s^{1-1/p}}\|B'\mathbf{v}\|_1 + \tau'\|A\mathbf{v}\| \qquad \text{for all } \mathbf{v} \in \mathbb{C}^n \text{ and all } S \subseteq [N] \text{ with } \operatorname{card}(S) = s. \tag{17}$$

Proof. Substituting the bound

$$\|B'\mathbf{v}\|_1 = \|(B'\mathbf{v})_S\|_1 + \|(B'\mathbf{v})_{\overline{S}}\|_1 \le s^{1-1/p}\|(B'\mathbf{v})_S\|_p + \|(B'\mathbf{v})_{\overline{S}}\|_1$$

into (17) and rearranging gives

$$\|(B'\mathbf{v})_S\|_p \le \frac{\rho'}{(1-\rho')s^{1-1/p}}\|(B'\mathbf{v})_{\overline{S}}\|_1 + \frac{\tau'}{1-\rho'}\|A\mathbf{v}\|.$$

We recognize the ℓ_p-robust null space property (9) with $\tau := \tau'/(1-\rho')$ and $\rho := \rho'/(1-\rho')$. The fact that $\rho < 1$ follows from $\rho' < 1/2$. $\square$

The second result is an observation that enables to control $\|(B^\dagger\mathbf{v})_S\|_2$ in terms of $\|B(B^\dagger\mathbf{v})_S\|_2$. It is useful in the upcoming arguments based on the restricted isometry property, since these are tailored to produce bounds on $\|B(B^\dagger\mathbf{v})_S\|_2$. It is essential here to enforce B' to coincide with the pseudo-inverse

$$B^\dagger := B^*(BB^*)^{-1}.$$

We recall that the smallest and largest singular values of B are simply denoted $\sigma_{\min}$ and $\sigma_{\max}$.

Lemma 7. For any $\mathbf{v} \in \mathbb{C}^n$, if S denotes an index set of s largest absolute entries of $B^\dagger \mathbf{v}$, then

$$\|(B^\dagger \mathbf{v})_S\|_2 \le \frac{\sqrt{s'/s}}{\sigma_{\min}} \|B(B^\dagger \mathbf{v})_S\|_2 + \frac{\sqrt{s/s'}}{2} \frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{s}} \qquad \text{for all integers } s' \ge s.$$

Proof. Let S' denote an index set of s' largest absolute entries of $B^\dagger \mathbf{v}$. We split $\|(B^\dagger \mathbf{v})_S\|_2^2$ as

$$\begin{aligned}
\|(B^\dagger \mathbf{v})_S\|_2^2 &= \langle B^\dagger \mathbf{v}, (B^\dagger \mathbf{v})_S \rangle = \langle B^\dagger \big(B(B^\dagger \mathbf{v})_{S'} + B(B^\dagger \mathbf{v})_{\overline{S'}} \big), (B^\dagger \mathbf{v})_S \rangle \\
&= \langle (BB^*)^{-1} B(B^\dagger \mathbf{v})_{S'}, B(B^\dagger \mathbf{v})_S \rangle + \langle B^\dagger B (B^\dagger \mathbf{v})_{\overline{S'}}, (B^\dagger \mathbf{v})_S \rangle \\
&\le \|(BB^*)^{-1} B\|_{2 \to 2} \|(B^\dagger \mathbf{v})_{S'}\|_2 \|B(B^\dagger \mathbf{v})_S\|_2 + \|B^\dagger B\|_{2\to 2} \|(B^\dagger \mathbf{v})_{\overline{S'}}\|_2 \|(B^\dagger \mathbf{v})_S\|_2.
\end{aligned}$$

We notice that the singular values of $(BB^*)^{-1}B$ are the inverses of the singular values of B, and that the nonzero singular values of $B^\dagger B$ are all equal to one, so that

$$\|(B^\dagger \mathbf{v})_S\|_2^2 \le \frac{1}{\sigma_{\min}} \|(B^\dagger \mathbf{v})_{S'}\|_2 \|B(B^\dagger \mathbf{v})_S\|_2 + \|(B^\dagger \mathbf{v})_{\overline{S'}}\|_2 \|(B^\dagger \mathbf{v})_S\|_2. \tag{18}$$

Because of the way S was chosen, we have

$$\frac{1}{s'} \|(B^\dagger \mathbf{v})_{S'}\|_2^2 \le \frac{1}{s} \|(B^\dagger \mathbf{v})_S\|_2^2, \qquad \text{so that} \quad \|(B^\dagger \mathbf{v})_{S'}\|_2 \le \sqrt{\frac{s'}{s}} \|(B^\dagger \mathbf{v})_S\|_2. \tag{19}$$

Because of the way S' was chosen (see [9, Lemma 7] or [6, Theorem 2]), we have

$$\|(B^\dagger \mathbf{v})_{\overline{S'}}\|_2 \le \frac{1}{2\sqrt{s'}} \|B^\dagger \mathbf{v}\|_1. \tag{20}$$

The result follows by substituting (19) and (20) into (18) and simplifying by $\|(B^\dagger \mathbf{v})_S\|_2$. $\square$

3.3 Proof of the Robust Null Space Property for Subgaussian Matrices

This subsection serves as a preamble to the next subsection by providing a version of Theorem 1 for subgaussian matrices instead of Weibull matrices. For tight frames and $p = 2$, this is exactly the result of [2]. In fact, we follow the article [2] by introducing the B-restricted isometry constant δ_s of order s for $A \in \mathbb{C}^{m \times n}$ as the smallest $\delta \ge 0$ such that

$$(1-\delta)\|B\mathbf{z}\|_2^2 \le \|AB\mathbf{z}\|_2^2 \le (1+\delta)\|B\mathbf{z}\|_2^2 \qquad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

As in the classical case $B = \mathrm{Id}$, it can be shown that, for any $\mathbf{z}, \mathbf{z}' \in \mathbb{C}^N$ such that $B\mathbf{z}$ and $B\mathbf{z}'$ are disjointly supported, we have

$$|\langle AB\mathbf{z}, AB\mathbf{z}' \rangle| \le \delta_t \|B\mathbf{z}\|_2 \|B\mathbf{z}'\|_2, \qquad t := \mathrm{card}(\mathrm{supp}(\mathbf{z}) \cup \mathrm{supp}(\mathbf{z}')).$$

As explained in [2], with high probability on the draw of a subgaussian matrix $A \in \mathbb{C}^{m \times n}$, the B-restricted isometry constant δ_{2s} can be made smaller than a prescribed threshold δ_* provided $m \ge c\, s \ln(eN/s)$, where the constant c depends on δ_*. A variant of Theorem 1 then follows from the fact that making the threshold δ_* small enough implies the validity of the estimates (6). When $\sigma_{\max} = \sigma_{\min}$ and $p = 2$, this was proved with $\delta_* = 0.08$ in [2] using cone and tube constraints formalism introduced in [3]. The argument below provides an alternative approach based on the robust null space property. As in [2], no serious attempt was made to reduce the threshold δ_* further. This could be achieved by adapting elaborate techniques used in the case $B = \mathrm{Id}$, such as the ones of [13].

Proposition 8. Given $B \in \mathbb{C}^{n \times N}$, if the B-restricted isometry constant δ_{2s} of a measurement matrix $A \in \mathbb{C}^{m \times n}$ satisfies

$$\delta_{2s} < \frac{\sigma_{\min}}{9\, \sigma_{\max}},$$

then the matrices A and $B^\dagger$ satisfy the ℓ_2-robust null space property of order s relative to the ℓ_2-norm on $\mathbb{C}^m$ and with constants $0 < \rho < 1$ and $\tau > 0$ depending only on δ_{2s}.

Proof. Let $\mathbf{v} \in \mathbb{C}^n$, and let $S = S_0$ denote an index set of s largest absolute entries of $B^\dagger \mathbf{v}$. Let us further introduce the index sets S_1 of next s largest absolute entries, S_2 of next s largest absolute entries, etc. It is classical to observe that

$$\|(B^\dagger \mathbf{v})_{S_k}\|_2 \leq \frac{\|(B^\dagger \mathbf{v})_{S_{k-1}}\|_1}{\sqrt{s}}, \quad k \geq 1,$$

so that a summation gives

$$\sum_{k \geq 1} \|(B^\dagger \mathbf{v})_{S_k}\|_2 \leq \frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{s}}. \tag{21}$$

We now use the properties of the B-restricted isometry constant to derive

$$\begin{aligned}
\|B(B^\dagger \mathbf{v})_{S_0}\|_2^2 &\leq \frac{1}{1-\delta_s} \|AB(B^\dagger \mathbf{v})_{S_0}\|_2^2 = \frac{1}{1-\delta_s} \langle AB(B^\dagger \mathbf{v})_{S_0}, ABB^\dagger \mathbf{v} - AB(B^\dagger \mathbf{v})_{\overline{S_0}} \rangle \\
&= \frac{1}{1-\delta_s} \langle AB(B^\dagger \mathbf{v})_{S_0}, A\mathbf{v} \rangle - \frac{1}{1-\delta_s} \sum_{k \geq 1} \langle AB(B^\dagger \mathbf{v})_{S_0}, AB(B^\dagger \mathbf{v})_{S_k} \rangle \\
&\leq \frac{1}{1-\delta_s} \|AB(B^\dagger \mathbf{v})_{S_0}\|_2 \|A\mathbf{v}\|_2 + \frac{1}{1-\delta_s} \sum_{k \geq 1} \delta_{2s} \|B(B^\dagger \mathbf{v})_{S_0}\|_2 \|B(B^\dagger \mathbf{v})_{S_k}\|_2 \\
&\leq \frac{\sqrt{1+\delta_s}}{1-\delta_s} \|B(B^\dagger \mathbf{v})_{S_0}\|_2 \|A\mathbf{v}\|_2 + \frac{\delta_{2s}}{1-\delta_s} \|B(B^\dagger \mathbf{v})_{S_0}\|_2 \sum_{k \geq 1} \sigma_{\max} \|(B^\dagger \mathbf{v})_{S_k}\|_2.
\end{aligned}$$

Using (21) and simplifying by $\|B(B^\dagger \mathbf{v})_{S_0}\|_2$ gives

$$\|B(B^\dagger \mathbf{v})_{S_0}\|_2 \leq \frac{\sqrt{1+\delta_s}}{1-\delta_s} \|A\mathbf{v}\|_2 + \frac{\delta_{2s}}{1-\delta_s} \sigma_{\max} \frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{s}}.$$

In view of Lemma 7 with $s' = 4s$, we finally obtain

$$\|(B^\dagger \mathbf{v})_S\|_2 \leq \frac{2\sqrt{1+\delta_s}}{\sigma_{\min}(1-\delta_s)} \|A\mathbf{v}\|_2 + \Big(\frac{2\,\sigma_{\max}}{\sigma_{\min}} \frac{\delta_{2s}}{1-\delta_s} + \frac{1}{4} \Big) \frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{s}}.$$

According to Lemma 6 and Theorem 5, it only remains to verify that

$$\rho' := \frac{2\sigma_{\max}}{\sigma_{\min}} \frac{\delta_{2s}}{1-\delta_s} + \frac{1}{4} < \frac{1}{2}.$$

In view of $\delta_s \leq \delta_{2s}$ and of $\sigma_{\max} \geq \sigma_{\min}$, this is guaranteed by $\delta_{2s} < \sigma_{\min}/(9\,\sigma_{\max})$. $\square$

3.4 Proof of the Robust Null Space Property for Weibull Matrices

The previous arguments do not cover Weibull matrices, since these matrices do not satisfy the restricted isometry property with the optimal number of measurements. Indeed, with $r = 1$ and $B = \mathrm{Id}$, it was observed in [1] that the restricted isometry property necessitates $m \geq c\,s \ln^2(eN/s)$. Nonetheless, the robust null space property extends to Weibull matrices, as proved in Proposition 10 below. The justification is based on the results of [8], which are applicable since a symmetric Weibull random matrix ξ is pregaussian in the sense that there exist constants $b, c > 0$ such that

$$\Pr(|\xi| \geq t) \leq b \exp(-ct), \qquad t \geq 0.$$

The key idea of [8] was to modify the restricted isometry property by replacing the inner norm by the ℓ_1-norm and by allowing the outer norm to depend on the probability distributions of the entries of A. The results, adapted to the present situation, are summarized in the lemma below. Actually, only the case $B = \mathrm{Id}$ was considered in [8], but the following generalization is simple enough to be left out — it duplicates the steps of [8], namely concentration inequality and covering arguments for (22), application of Khintchine's inequality (see e.g. [11]) for (24).

Lemma 9. Let $B \in \mathbb{C}^{n\times N}$ and $\delta_* > 0$ be given. If $A \in \mathbb{R}^{m\times n}$ is a Weibull random matrix with exponent $r \geq 1$, then there exist constants $c_1, c_2 > 0$ depending only on δ_* such that the property

$$(22) \qquad (1-\delta_*) \,/\!\!/\, B\mathbf{z} /\!\!/ \leq \|AB\mathbf{z}\|_1 \leq (1+\delta_*) \,/\!\!/\, B\mathbf{z} /\!\!/, \qquad \text{all } s\text{-sparse } \mathbf{z} \in \mathbb{C}^N,$$

holds with probability at least $1 - 2\exp(-c_1 m)$, as long as $m \geq c_2\, s \ln(eN/s)$. The outer norm is defined by

$$(23) \qquad /\!\!/\, \mathbf{v} /\!\!/ := m \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left| \sum_{j=1}^{n} t_j v_j \right| f_r(t_1) \cdots f_r(t_n)\, dt_1 \cdots dt_n, \qquad \mathbf{v} \in \mathbb{C}^n,$$

where f_r denotes the probability density function of the entries of A. This norm is comparable to the Euclidean norm: if μ_r and σ_r denote the first absolute moment and the variance of f_r, then, for all $\mathbf{v} \in \mathbb{C}^n$,

$$(24) \qquad \frac{\mu_r\, m}{\sqrt{8}} \|\mathbf{v}\|_2 \leq /\!\!/\, \mathbf{v} /\!\!/ \leq \sigma_r\, m \|\mathbf{v}\|_2, \qquad \text{hence} \qquad \frac{\sqrt{m}}{4} \|\mathbf{v}\|_2 \leq /\!\!/\, \mathbf{v} /\!\!/ \leq \sqrt{m}\, \|\mathbf{v}\|_2.$$

With this lemma at hand, we can now state and prove the main result of this subsection. The proof follows the arguments of [8, Theorem 5.1], which established (in the case $B = \mathrm{Id}$) that the modified restricted isometry property (22) and the equivalence of norms (24) together imply the ℓ_2-robust null space property — in fact, only the ideal case of exactly sparse vectors measured with perfect accuracy was formally treated, while stability and robustness were stated without proof there.

Proposition 10. Let $B \in \mathbb{C}^{n \times N}$, $n \leq N$, be a matrix with smallest and largest singular values $\sigma_{\min}$ and $\sigma_{\max}$. If $A \in \mathbb{C}^{m \times n}$ is a Weibull random matrix with exponent $r \geq 1$, then with probability at least $1 - 2\exp(-c_1 m)$, the matrices A and $B^\dagger$ satisfy the ℓ_2-robust null space property of order s with constants $0 < \rho < 1$ and $\tau > 0$, as long as $m \geq c_2\, s \ln(eN/s)$. The constants ρ and c_1 are universal, while τ depends on $\sigma_{\min}$ and c_2 depends on $\sigma_{\max}/\sigma_{\min}$.

Proof. Let us set

$$t := \Big(\frac{15\,\sigma_{\max}}{\sigma_{\min}}\Big)^2 s \qquad \text{and} \qquad \delta_* := \frac{1}{9}.$$

Under the hypotheses of Proposition 10, Lemma 9 guarantees that

$$(1-\delta_*) \mathbin{/\!\!/} B\mathbf{z} \mathbin{/\!\!/} \leq \|AB\mathbf{z}\|_1 \leq (1+\delta_*) \mathbin{/\!\!/} B\mathbf{z} \mathbin{/\!\!/} \qquad \text{for all } t\text{-sparse } \mathbf{z} \in \mathbb{C}^N. \tag{25}$$

Let us now consider a vector $\mathbf{v} \in \mathbb{C}^N$, and let S denote an index set of s largest absolute entries of $B^\dagger\mathbf{v}$. Let us further introduce the index sets $T_0 \supseteq S$ of t largest absolute entries of $B^\dagger\mathbf{v}$, T_1 of next t largest absolute entries, etc. We observe that

$$\sum_{k \geq 1} \|(B^\dagger\mathbf{v})_{T_k}\|_2 \leq \frac{\|B^\dagger\mathbf{v}\|_1}{\sqrt{t}}. \tag{26}$$

We use properties (25), (24), and (26) to derive

$$\begin{aligned}
\mathbin{/\!\!/} B(B^\dagger\mathbf{v})_{T_0} \mathbin{/\!\!/} &\leq \frac{1}{1-\delta_*}\|AB(B^\dagger\mathbf{v})_{T_0}\|_1 = \frac{1}{1-\delta_*}\|ABB^\dagger\mathbf{v} - AB(B^\dagger\mathbf{v})_{\overline{T_0}}\|_1 \\
&\leq \frac{1}{1-\delta_*}\Big[\|A\mathbf{v}\|_1 + \sum_{k\geq 1}\|AB(B^\dagger\mathbf{v})_{T_k}\|_1\Big] \\
&\leq \frac{1}{1-\delta_*}\Big[\|A\mathbf{v}\|_1 + \sum_{k\geq 1}(1+\delta_*) \mathbin{/\!\!/} B(B^\dagger\mathbf{v})_{T_k} \mathbin{/\!\!/}\Big] \\
&\leq \frac{1}{1-\delta_*}\Big[\|A\mathbf{v}\|_1 + (1+\delta_*)\sum_{k\geq 1}\sqrt{m}\|B(B^\dagger\mathbf{v})_{T_k}\|_2\Big] \\
&\leq \frac{1}{1-\delta_*}\Big[\|A\mathbf{v}\|_1 + (1+\delta_*)\sqrt{m}\sum_{k\geq 1}\sigma_{\max}\|(B^\dagger\mathbf{v})_{T_k}\|_2\Big] \\
&\leq \frac{1}{1-\delta_*}\Big[\|A\mathbf{v}\|_1 + (1+\delta_*)\sqrt{m}\,\sigma_{\max}\frac{\|B^\dagger\mathbf{v}\|_1}{\sqrt{t}}\Big].
\end{aligned}$$

Using (24) once again, we arrive at

$$\|B(B^\dagger \mathbf{v})_{T_0}\|_2 \le \frac{4}{1-\delta_*}\frac{\|A\mathbf{v}\|_1}{\sqrt{m}} + \frac{4(1+\delta_*)}{1-\delta_*}\sigma_{\max}\frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{t}}.$$

In view of Lemma 7 with t taking the role of both s and s', we obtain

$$\begin{aligned}\|(B^\dagger \mathbf{v})_{T_0}\|_2 &\le \frac{1}{\sigma_{\min}}\|B(B^\dagger \mathbf{v})_{T_0}\|_2 + \frac{1}{2}\frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{t}} \\ &\le \frac{4}{\sigma_{\min}(1-\delta_*)}\frac{\|A\mathbf{v}\|_1}{\sqrt{m}} + \Big(\frac{4(1+\delta_*)}{1-\delta_*}\sqrt{\frac{s}{t}}\frac{\sigma_{\max}}{\sigma_{\min}} + \frac{1}{2}\sqrt{\frac{s}{t}}\Big)\frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{s}}.\end{aligned}$$

Substituting the values of t and δ_*, while using $\|(B^\dagger \mathbf{v})_S\|_2 \le \|(B^\dagger \mathbf{v})_{T_0}\|_2$, we conclude

$$\|(B^\dagger \mathbf{v})_S\|_2 \le \rho'\frac{\|B^\dagger \mathbf{v}\|_1}{\sqrt{s}} + \tau'\frac{\|A\mathbf{v}\|_1}{\sqrt{m}} \qquad \text{with } \rho' \approx 1/3 < 1/2 \text{ and } \tau' = 4.5/\sigma_{\min}.$$

Taking $\|A\mathbf{v}\|_1/\sqrt{m} \le \|A\mathbf{v}\|_2$ into account, Lemma 6 finally implies the ℓ_2-robust null space property relative to the ℓ_2-norm. □

Remark. The proof actually shows the stronger statement that Weibull matrices satisfy the ℓ_2-robust null space property relative to the norm $\|\cdot\|_1/\sqrt{m}$.

Remark. These arguments would also apply to subgaussian matrices. The arguments of the previous subsection and of the present subsection would both lead to sufficient conditions in terms of restricted isometry constants that read $\delta_t < \delta_*$ with

$$t = c_1\, s, \quad \delta_* = c_2\,\frac{\sigma_{\max}}{\sigma_{\min}}, \qquad \text{and} \qquad t = c_1\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2 s, \quad \delta_* = c_2,$$

respectively. Both conditions are met when $m/\ln(eN/m) \ge c(\delta_*)\, t$. It is possible to show that $c(\delta_*) = c_3/\delta_*^2$ in the subgaussian case, while the same behavior is expected in the Weibull case. If so, the two sufficient conditions are not essentially different, since they both require the number of measurement to satisfy

$$\frac{m}{\ln(eN/m)} \approx \frac{c_1\, c_3}{c_2^2}\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)^2 s.$$

4 Equality-Constrained ℓ_1-Minimization

This section is devoted to the proof of Theorem 2. The deterministic arguments showing that the estimate (7) is a consequence of the robust null space property and the quotient property combined are presented in Subsection 4.1. The probabilistic arguments showing that Weibull matrices satisfy the quotient property are presented in Subsection 4.3, after some auxiliary lemmas are collected in Subsection 4.2.

4.1 Sufficiency of the Robust Null Space and Quotient Properties

The result below outlines the relevance of the ℓ_1-quotient property when generalizing estimates (5) to random matrices that are not Gaussian. The arguments are essentially contained in [15], although they were not isolated because the restricted isometry property was assumed throughout. For our purpose, it is crucial to replace this assumption by the weaker robust null space property, since Weibull matrices do not satisfy the restricted isometry property in the optimal regime $m \asymp s \ln(eN/s)$.

Theorem 11. Given $q \geq 1$, suppose that the matrix $A \in \mathbb{C}^{m \times N}$ satisfies the ℓ_1-quotient property with constant $d > 0$ relative to a norm $\|\cdot\|$ on $\mathbb{C}^m$, as well as the ℓ_q-robust null space property of order $s = c\, s_*$ with constants $0 < \rho < 1$ and $\tau > 0$ relative to the norm $s_*^{1/q-1/2}\|\cdot\|$. Then, for any $1 \leq p \leq q$, the bounds

$$\|\mathbf{x} - \Delta_1(A\mathbf{x} + \mathbf{e})\|_p \leq \frac{C}{s^{1-1/p}}\, \sigma_s(\mathbf{x})_1 + D\, s_*^{1/p-1/2}\, \|\mathbf{e}\|$$

hold for all $\mathbf{x} \in \mathbb{C}^N$ and all $\mathbf{e} \in \mathbb{C}^m$. The constants $C, D > 0$ depend only on ρ, τ, c, and d.

This theorem is a consequence of the following two lemmas and of the observation that the ℓ_q-robust null space property of order $s = c\, s_*$ relative to $s_*^{1/q-1/2}\|\cdot\|$ implies the ℓ_p-robust null space property of order $s = c\, s_*$ relative to $s_*^{1/p-1/2}\|\cdot\|$ for any $p \leq q$.

Lemma 12. Given $q \geq 1$, suppose that $A \in \mathbb{C}^{m \times N}$ satisfies the ℓ_1-quotient property with constant $d > 0$ relative to a norm $\|\cdot\|$, as well as the ℓ_q-robust null space property of order $s = c\, s_*$ with constant $\rho > 0$ and $\tau > 0$ relative to the norm $s_*^{1/q-1/2}\|\cdot\|$. Then A also satisfies the simultaneous (ℓ_1, ℓ_q)-quotient property with constants $d, d' > 0$ relative to $\|\cdot\|$, in the sense that

$$\text{for all } \mathbf{e} \in \mathbb{C}^m, \text{ there exists } \mathbf{v} \in \mathbb{C}^N \text{ with } A\mathbf{v} = \mathbf{e} \text{ and } \begin{cases} \|\mathbf{v}\|_q \leq d'\, s_*^{1/q-1/2}\, \|\mathbf{e}\|, \\ \|\mathbf{v}\|_1 \leq d\, s_*^{1/2}\, \|\mathbf{e}\|. \end{cases}$$

Proof. Let us consider a vector $\mathbf{e} \in \mathbb{C}^m$. By the ℓ_1-quotient property, there exists $\mathbf{v} \in \mathbb{C}^N$ such that $A\mathbf{v} = \mathbf{e}$ and $\|\mathbf{v}\|_1 \leq d\, s_*^{1/2}\, \|\mathbf{e}\|$. We shall prove that $\|\mathbf{v}\|_q \leq d'\, s_*^{1/q-1/2}\, \|\mathbf{e}\|$ for some constant $d' > 0$. For an index set S of s largest absolute entries of $\mathbf{v}$, we have

$$\|\mathbf{v}_{\overline{S}}\|_q = \sigma_s(\mathbf{v})_q \leq \frac{1}{s^{1-1/q}} \|\mathbf{v}\|_1.$$

Moreover, in view of the ℓ_q-robust null space property of order s, we have

$$\|\mathbf{v}_S\|_q \leq \frac{\rho}{s^{1-1/q}}\, \|\mathbf{v}_{\overline{S}}\|_1 + \tau s^{1/q-1/2}\|A\mathbf{v}\| \leq \frac{\rho}{s^{1-1/q}}\, \|\mathbf{v}\|_1 + \tau' s_*^{1/q-1/2}\|\mathbf{e}\|.$$

It follows that

$$\|\mathbf{v}\|_q \leq \|\mathbf{v}_{\overline{S}}\|_q + \|\mathbf{v}_S\|_q \leq \frac{1+\rho}{s^{1-1/q}}\|\mathbf{v}\|_1 + \tau' s_*^{1/q-1/2}\|\mathbf{e}\|.$$

The estimate $\|\mathbf{v}\|_1 \leq d\, s_*^{1/2}\, \|\mathbf{e}\|$ finally yields the required result. $\square$

The next lemma involves the notion of instance optimality, recalled in (16), taken for $B' = \mathrm{Id}$.

Lemma 13. Given $q \geq 1$, suppose that the matrices $A \in \mathbb{C}^{m \times N}$ satisfy the simultaneous (ℓ_1, ℓ_q)-quotient property relative to a norm $\|\cdot\|$, and that a reconstruction map $\Delta : \mathbb{C}^m \to \mathbb{C}^N$ makes the pair (A, Δ) mixed (ℓ_q, ℓ_1)-instance optimal of order $c\, s_*$ with constant $C > 0$. Then, for $s \leq c\, s^*$, the bound

$$\|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_q \leq \frac{C}{s^{1-1/q}}\, \sigma_s(\mathbf{x})_1 + D\, s_*^{1/q-1/2}\, \|\mathbf{e}\|, \qquad D = Cc^{-1/2}d + c^{1/2-1/q}d',$$

holds for all $\mathbf{x} \in \mathbb{C}^N$ and all $\mathbf{e} \in \mathbb{C}^m$.

Proof. For $\mathbf{e} \in \mathbb{C}^m$, the simultaneous (ℓ_1, ℓ_q)-quotient property yields the existence of $\mathbf{v} \in \mathbb{C}^N$ such that $A\mathbf{v} = \mathbf{e}$ and

$$\|\mathbf{v}\|_q \leq d'\, s_*^{1/q-1/2}\, \|\mathbf{e}\|, \qquad \|\mathbf{v}\|_1 \leq d\, s_*^{1/2}\, \|\mathbf{e}\|. \tag{27}$$

Next, for $\mathbf{x} \in \mathbb{C}^N$, the instance optimality of order $s = c\, s_*$ yields

$$\begin{aligned}\|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_q &= \|\mathbf{x} - \Delta(A(\mathbf{x} + \mathbf{v}))\|_q \leq \|\mathbf{x} + \mathbf{v} - \Delta(A(\mathbf{x} + \mathbf{v}))\|_q + \|\mathbf{v}\|_q \\ &\leq \frac{C}{s^{1-1/q}}\, \sigma_s(\mathbf{x} + \mathbf{v})_1 + \|\mathbf{v}\|_q \leq \frac{C}{s^{1-1/q}} \big(\sigma_s(\mathbf{x})_1 + \|\mathbf{v}\|_1\big) + \|\mathbf{v}\|_q.\end{aligned}$$

Substituting (27) into the latter gives the required result for $s = c\, s_*$, and the result for $s \leq c\, s_*$ follows by monotonicity. □

4.2 Auxiliary Lemmas

In order to derive Theorem 2 from Theorem 11, it remains to prove that Weibull random matrices satisfy the ℓ_1-quotient property, since the ℓ_2-robust null space property was already established in Proposition 10. The proof requires some preliminary results presented in this section. The first result is a rephrasing of the quotient property in a more amenable form. This observation was implicitly used in the two articles [15, 5] concerning stability and robustness of the equality-constrained ℓ_1-minimization. It only differs in the fact that the norm on $\mathbb{C}^N$ is not specified here. A proof is included for completeness.

Lemma 14. Given $p \geq 1$, the matrices $A \in \mathbb{C}^{m \times N}$ satisfies the ℓ_p-quotient property with constant d relative to a norm $\|\cdot\|$ on $\mathbb{C}^m$ if and only if

$$\|\mathbf{u}\|_* := \max_{\|\mathbf{e}\|=1} |\langle \mathbf{e}, \mathbf{u} \rangle| \leq d\, s_*^{1/p-1/2} \|A^*\mathbf{u}\|_{p^*} \qquad \text{for all } \mathbf{u} \in \mathbb{C}^m, \tag{28}$$

where $s_* = m/\ln(eN/m)$ and where $p^* := p/(p-1)$ denotes the conjugate exponent of p.

Proof. Suppose first that the ℓ_p-quotient property holds. We consider $\mathbf{u} \in \mathbb{C}^m$, and we write $\|\mathbf{u}\|_* = |\langle \mathbf{e}, \mathbf{u} \rangle|$ with $\|\mathbf{e}\| = 1$. By the ℓ_p-quotient property, there exists $\mathbf{v} \in \mathbb{C}^N$ such that $A\mathbf{v} = \mathbf{e}$ and $\|\mathbf{v}\|_p \le d\, s_*^{1/p-1/2}$. We then derive

$$\|\mathbf{u}\|_* = |\langle \mathbf{e}, \mathbf{u} \rangle| = |\langle A\mathbf{v}, \mathbf{u} \rangle| = |\langle \mathbf{v}, A^*\mathbf{u} \rangle| \le \|\mathbf{v}\|_p \|A^*\mathbf{u}\|_{p^*} \le d\, s_*^{1/p-1/2} \|A^*\mathbf{u}\|_{p^*}.$$

This shows the validity of (28).
Suppose conversely that (28) holds. We assume first that $p > 1$. For a nonzero $\mathbf{e} \in \mathbb{C}^m$, let $\mathbf{v} \in \mathbb{C}^N$ be a minimizer of $\|\mathbf{z}\|_p$ subject to $A\mathbf{z} = \mathbf{e}$.[3] Our goal is to show that $\|\mathbf{v}\|_p \le d s_*^{1/p-1/2} \|\mathbf{e}\|$. We fix a vector $\mathbf{u} \in \ker A$ for a while. Given $\tau = t e^{-i\theta}$ with $t > 0$ small enough to have $\mathbf{v} + \tau \mathbf{u} \neq 0$, we consider the vector $\mathbf{w}^\tau \in \mathbb{C}^N$ whose entries are given by

$$w_j^\tau := \frac{\operatorname{sgn}(u_j + \tau v_j)\, |u_j + \tau v_j|^{p-1}}{\|\mathbf{v} + \tau \mathbf{u}\|_p^{p-1}}, \qquad j \in [N].$$

We notice that $\langle \mathbf{w}^\tau, \mathbf{v} + \tau \mathbf{u} \rangle = \|\mathbf{v} + \tau \mathbf{u}\|_p$ and that $\|\mathbf{w}^\tau\|_{p^*} = 1$. We also notice that the vector $\mathbf{w} := \lim_{\tau \to 0} \mathbf{w}^\tau$ is well-defined and independent of $\mathbf{u}$, thanks to the assumption $p > 1$. It satisfies $\langle \mathbf{w}, \mathbf{v} \rangle = \|\mathbf{v}\|_p$ with $\|\mathbf{w}\|_{p^*} = 1$. The definition of $\mathbf{v}$ yields

$$\Re \langle \mathbf{w}^\tau, \mathbf{v} \rangle \le \|\mathbf{v}\|_p \le \|\mathbf{v} + \tau \mathbf{u}\|_p = \Re \langle \mathbf{w}^\tau, \mathbf{v} + \tau \mathbf{u} \rangle,$$

so that $\Re \langle \mathbf{w}^\tau, e^{-i\theta} \mathbf{u} \rangle \ge 0$. Taking the limit as t tends to zero, we obtain independently of θ that $e^{i\theta} \Re \langle \mathbf{w}, \mathbf{u} \rangle \ge 0$, hence $\langle \mathbf{w}, \mathbf{u} \rangle = 0$. Since this is true for all $\mathbf{u} \in \ker A$, we deduce that $\mathbf{w} \in (\ker A)^\perp = \operatorname{im} A^*$. Therefore we can write $\mathbf{w} = A^* \mathbf{y}$ for some $\mathbf{y} \in \mathbb{C}^m$. According to (28), we have $\|\mathbf{y}\|_* \le d s_*^{1/p-1/2}$. It follows that

$$\|\mathbf{v}\|_p = \langle \mathbf{w}, \mathbf{v} \rangle = \langle A^* \mathbf{y}, \mathbf{v} \rangle = \langle \mathbf{y}, A\mathbf{v} \rangle = \langle \mathbf{y}, \mathbf{e} \rangle \le \|\mathbf{y}\|_* \|\mathbf{e}\| \le d s_*^{1/p-1/2} \|\mathbf{e}\|.$$

This establishes the ℓ_p-quotient property in the case $p > 1$. We use some limiting arguments for the case $p = 1$. Precisely, we consider a sequence of numbers $p_n > 1$ converging to 1. For each n, in view of $\|A^*\mathbf{e}\|_\infty \le \|A^*\mathbf{e}\|_{p_n^*}$, the property (28) for $p = 1$ implies a similar property for $p = p_n$ provided d is changed to $d' := d s_*^{1/p_n^*}$. Given $\mathbf{e} \in \mathbb{C}^m$, the previous argument yields a vector $\mathbf{v}^n \in \mathbb{C}^N$ with $A\mathbf{v}^n = \mathbf{e}$ and $\|\mathbf{v}^n\|_{p_n} \le d' s_*^{1/p_n - 1/2} \|\mathbf{e}\| = d s_*^{1/2} \|\mathbf{e}\|$. Since the sequence $(\mathbf{v}^n)$ is bounded in ℓ_∞-norm, it has a convergent subsequence. Denoting by $\mathbf{v} \in \mathbb{C}^N$ its limit, we obtain $A\mathbf{v} = \mathbf{e}$ and $\|\mathbf{v}\|_1 \le d s_*^{1/2} \|\mathbf{e}\|$ by letting n tend to infinity. This settles the case $p = 1$. $\square$

Remark. For a real matrix $A \in \mathbb{R}^{m \times N}$, the property (28) is equivalent, up to modification of the constant d, to the real version

$$\|\mathbf{u}\|_* \le d\, s^{1/p-1/2} \|A^*\mathbf{u}\|_{p^*} \qquad \text{for all } \mathbf{u} \in \mathbb{R}^m.$$

As a matter of fact, it is this real version of the quotient property that is established below.

[3] One can consider such a minimizer, since (28) implies $\ker A^* = \{0\}$, i.e., $\operatorname{im} A = (\ker A^*)^\perp = \mathbb{C}^m$.

Next, we review some results from the article [10] of Gluskin and Kwapień, which deals with sums of independent random variables with logarithmically concave tails. To express these results, we need the norm defined, for $r \geq 1$, for an integer $k \geq 1$, and for a vector $\mathbf{u} \in \mathbb{R}^N$ by

$$\|\mathbf{u}\|_{(k)} := \max\left\{\gamma_r k^{1/r}\|\mathbf{u}_k\|_{r^*}, k^{1/2}\|\mathbf{u}-\mathbf{u}_k\|_2\right\}, \qquad \gamma_r := \Big(\frac{1}{r}\Big)^{1/r}\Big(\frac{1}{r^*}\Big)^{1/r^*}, \quad r^* := \frac{r}{r-1},$$

where $\mathbf{u}_k$ denotes a best k-term approximation to $\mathbf{u}$. It is in fact the quantity

$$F(t,\mathbf{u}) := \max\{k \geq 1 : \|\mathbf{u}\|_{(k)} \leq t\}$$

that is involved in the following lemma.

Lemma 15. Let $(\xi_i)_{i=1}^\infty$ be a sequence of independent symmetric Weibull random variables with exponent $r \geq 1$ satisfying

$$\Pr(|\xi_i| \geq t) = \exp(-t^r), \qquad t \geq 0.$$

There exist constants $\alpha_r, \beta_r > 0$ such that, for any $\mathbf{u} \in \mathbb{R}^N$ and any $t \geq \|\mathbf{u}\|_{(1)}$,

$$\exp(-F(\alpha_r t,\mathbf{u})) \leq \Pr\Big(\Big|\sum_{i=1}^\infty u_i\xi_i\Big| > t\Big) \leq \exp(-F(\beta_r t,\mathbf{u})).$$

We shall only use a consequence of the lower bound. For $1 \leq r \leq 2$, we remark that $\gamma_r \geq 1/2$ and that $\|\mathbf{u}_k\|_2 \leq k^{1/2-1/r^*}\|\mathbf{u}_k\|_{r^*}$. Thus, for any $\mathbf{u} \in \mathbb{R}^N$,

$$\|\mathbf{u}\|_{(k)} \geq \frac{1}{2}\max\left\{k^{1/2}\|\mathbf{u}_k\|_2, k^{1/2}\|\mathbf{u}-\mathbf{u}_k\|_2\right\} \geq \frac{k^{1/2}}{2\sqrt{2}}\big(\|\mathbf{u}_k\|_2^2 + \|\mathbf{u}-\mathbf{u}_k\|_2^2\big)^{1/2} = \frac{k^{1/2}}{2\sqrt{2}}\|\mathbf{u}\|_2.$$

It follows that

$$F(t,\mathbf{u}) \leq \max\left\{k \geq 1 : \frac{k^{1/2}}{2\sqrt{2}}\|\mathbf{u}\|_2 \leq t\right\} \leq \frac{8t^2}{\|\mathbf{u}\|_2^2}.$$

Therefore, for $1 \leq r \leq 2$, the lower bound yields the following estimate, valid for $t \geq \|\mathbf{u}\|_2$,

$$\Pr\Big(\Big|\sum_{i=1}^\infty u_i\xi_i\Big| > t\Big) \geq \exp\Big(-\frac{8\alpha_r^2 t^2}{\|\mathbf{u}\|_2^2}\Big). \tag{29}$$

The final auxiliary result, although not explicitly stated, appears in [8] as an intermediate step towards the derivation of Lemma 9. It is formulated for Weibull random matrices below, but it holds for pregaussian random matrices in general.

Lemma 16. Let $\delta_* > 0$ be given. If $A' \in \mathbb{R}^{m\times s}$ is a Weibull random matrix with exponent $r \geq 1$ and if $/\!\!/ \cdot /\!\!/$ is the norm defined in (23), then there exist constants $c_1, c_2 > 0$ such that the property

$$(1-\delta_*)\, /\!\!/ \mathbf{z} /\!\!/ \leq \|A'\mathbf{z}\|_1 \leq (1+\delta_*)\, /\!\!/ \mathbf{z} /\!\!/, \qquad \text{all } \mathbf{z} \in \mathbb{C}^s,$$

holds with probability at least $1 - 2\exp(-c_1 m)$, as long as $m \geq c_2\, s$.

4.3 Proof of the Quotient Property for Weibull Matrices

We are now in a position to prove the ℓ_1-quotient property for Weibull random matrices. The following proposition, combined with Proposition 10 and Theorem 11, implies Theorem 2.

Proposition 17. A Weibull random matrix $A \in \mathbb{R}^{m \times N}$ with exponent $1 \leq r \leq 2$ satisfies the ℓ_1-quotient property with constant $d = 16\alpha_r/\sqrt{\Gamma(1+2/r)}$ relative to the ℓ_2-norm with probability at least $1 - 3\exp(-C_1\sqrt{mN})$ provided $N \geq C_2 m$. The constant $C_1 > 0$ is universal, while the constant $C_2 > 0$ depends on α_r.

Remark. According to Remark 4.2, we have to prove that $\|\mathbf{e}\|_2 \leq d\sqrt{s_*}\|A^*\mathbf{e}\|_\infty$ for all $\mathbf{e} \in \mathbb{R}^m$. We shall prove below the stronger result that $\|\mathbf{e}\|_2 \leq d\sqrt{s_*}|||A^*\mathbf{e}|||$ for the norm defined on $\mathbb{R}^N$ by

$$|||\mathbf{z}||| := \frac{1}{2h}\sum_{\ell=1}^{2h}\|\mathbf{z}_{T_\ell}\|_\infty$$

relatively to a fixed partition $T_1, \ldots, T_{2h}$ of $\{1, \ldots, N\}$. The integer h is chosen so that $h \leq N/4$, and the partition $T_1, \ldots, T_{2h}$ is chosen so that the size t_ℓ of each T_ℓ satisfies

$$\left\lfloor \frac{N}{2h} \right\rfloor \leq t_\ell \leq 2\left\lfloor \frac{N}{2h} \right\rfloor, \qquad \text{hence} \quad \frac{N}{3h} \leq t_\ell \leq \frac{N}{h}.$$

Besides the property $|||\mathbf{z}||| \leq \|\mathbf{z}\|_\infty$, another key feature of the above norm is the existence, for any $\mathbf{z} \in \mathbb{R}^N$, of a subset H of $\{1, \ldots, 2h\}$ of cardinality h such that

$$\|\mathbf{z}_{T_\ell}\|_\infty \leq 2|||\mathbf{z}||| \qquad \text{for all } \ell \in H.$$

The introduction of this norm, taken from [14], is essential to reduce the cardinality of the ε-net in (30) below. It allows to relax the condition $N \geq m\ln^2(m)$ found in [15] for Gaussian matrices. It also allows to correct the arguments of [5] for Bernoulli matrices. There, an ε-net was chosen relative to a norm depending on the random matrix, so calling upon probability estimates valid for fixed vectors do not apply to the elements of such an ε-net.

Proof. We want to prove that, for all $\mathbf{e} \in \mathbb{R}^m$,

$$\|\mathbf{e}\|_2 \leq d\sqrt{s_*}|||A^*\mathbf{e}|||.$$

We start by considering individual vectors $\mathbf{e} \in \mathbb{R}^m$, for which we show that the probabilities

$\Pr(\|\mathbf{e}\|_2 > d'\sqrt{s_*}|\!|\!|A^*\mathbf{e}|\!|\!|)$, $d' = d/2$, are exponentially small. Fixing $\mathbf{e} \in \mathbb{R}^m$, we first observe that

$$\begin{aligned}
\Pr\left(\|\mathbf{e}\|_2 > d'\sqrt{s_*}|\!|\!|A^*\mathbf{e}|\!|\!|\right) &\le \Pr\left(\|(A^*\mathbf{e})_{T_\ell}\|_\infty < \frac{2\|\mathbf{e}\|_2}{d'\sqrt{s_*}} \text{ for all } \ell \text{ in some } H \subseteq [2h], \operatorname{card}(H) = h \right) \\
&\le \sum_{H\subseteq[2h],\operatorname{card}(H)=h} \Pr\left(\max_{j\in T_\ell} |(A^*\mathbf{e})_j| < \frac{2\|\mathbf{e}\|_2}{d'\sqrt{s_*}} \text{ for all } \ell \in H \right) \\
&= \sum_{H\subseteq[2h],\operatorname{card}(H)=h} \Pr\left(|(A^*\mathbf{e})_j| < \frac{2\|\mathbf{e}\|_2}{d'\sqrt{s_*}} \text{ for all } j \in \cup_{\ell\in H} T_\ell \right) \\
&= \sum_{H\subseteq[2h],\operatorname{card}(H)=h} \prod_{j\in\cup_{\ell\in H}T_\ell} \Pr\left(|(A^*\mathbf{e})_j| < \frac{2\|\mathbf{e}\|_2}{d'\sqrt{s_*}} \right).
\end{aligned}$$

The last equality follows from the independence of the $(A^*\mathbf{e})_j$, $j \in \cup_{\ell\in H} T_\ell$, which are given by

$$(A^*\mathbf{e})_j = \sum_{i=1}^m a_{i,j} e_i = \sum_{i=1}^m \frac{\sqrt{\Gamma(1+2/r)}\, e_i}{\sqrt{m}} \xi_{i,j}, \qquad \xi_{i,j} := \frac{\sqrt{m}}{\sqrt{\Gamma(1+2/r)}} a_{i,j}.$$

Since $N \ge C_2 m$, choosing the constant C_2 large enough makes an application of (29) possible. We obtain

$$\begin{aligned}
\Pr\left(\left|(A^*\mathbf{e})_j\right| < \frac{2\|\mathbf{e}\|_2}{d'\sqrt{s_*}} \right) &= 1 - \Pr\left(\left| \sum_{i=1}^m \frac{\sqrt{\Gamma(1+2/r)}\, e_i}{\sqrt{m}} \xi_{i,j} \right| \ge \frac{2\|\mathbf{e}\|_2}{d'\sqrt{s_*}} \right) \\
&\le 1 - \exp\left(- \frac{32\alpha_r^2 m \|\mathbf{e}\|_2^2}{d'^2 s_* \Gamma(1+2/r) \|\mathbf{e}\|_2^2} \right) = 1 - \exp\left(- \frac{32\alpha_r^2 \ln(eN/m)}{d'^2 \Gamma(1+2/r)} \right) \\
&= 1 - \left(\frac{m}{eN}\right)^{1/2} \le \exp\left(- \sqrt{\frac{m}{eN}} \right),
\end{aligned}$$

where we have made use of the value $d' = 8\alpha_r/\sqrt{\Gamma(1+2/r)}$. It follows that

$$\begin{aligned}
\Pr\left(\|\mathbf{e}\|_2 > d'\sqrt{s_*}|\!|\!|A^*\mathbf{e}|\!|\!|\right) &\le \sum_{H\subseteq[2h],\operatorname{card}(H)=h} \exp\left(- \sqrt{\frac{m}{eN}} \right)^{\operatorname{card}(\cup_{\ell\in H}T_\ell)} \le \binom{2h}{h} \exp\left(- \sqrt{\frac{m}{eN}} \right)^{N/3} \\
&= \binom{2h}{h} \exp\left(- \frac{\sqrt{mN}}{3\sqrt{e}} \right).
\end{aligned}$$

Next, we use covering arguments. For $0 < \varepsilon < 1$ to be chosen later, we introduce an ε-covering $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ of the unit sphere of ℓ_2^m relative to ℓ_2 with cardinality $k \le (1+2/\varepsilon)^m \le (3/\varepsilon)^m$, see e.g. [12, Chapter 3]. Let us suppose that there exists $\mathbf{e} \in \mathbb{R}^m$ with $\|\mathbf{e}\|_2 > d\sqrt{s_*}|\!|\!|A^*\mathbf{e}|\!|\!|$. Without loss of generality, we assume that $\|\mathbf{e}\|_2 = 1$, so that $\|\mathbf{e} - \mathbf{e}_i\|_2 \le \varepsilon$ for some $i \in [k]$. We then have

$$\begin{aligned}
d\sqrt{s_*}|\!|\!|A^*\mathbf{e}_i|\!|\!| &\le d\sqrt{s_*}|\!|\!|A^*\mathbf{e}|\!|\!| + d\sqrt{s_*}|\!|\!|A^*(\mathbf{e}-\mathbf{e}_i)|\!|\!| \le \|\mathbf{e}\|_2 + \frac{d\sqrt{s_*}}{2h} \sum_{\ell=1}^{2h} \|(A^*(\mathbf{e}-\mathbf{e}_i))_{T_\ell}\|_\infty \\
&\le 1 + \frac{d\sqrt{s_*}}{2h} \sum_{\ell=1}^{2h} \|(A^*(\mathbf{e}-\mathbf{e}_i))_{T_\ell}\|_1 = 1 + \frac{d\sqrt{s_*}}{2h} \|A^*(\mathbf{e}-\mathbf{e}_i)\|_1.
\end{aligned}$$

Since $N \geq C_2 m$, choosing the constant C_2 large enough makes it possible to apply Lemma 16 for $\delta_* = 1$ and for the renormalized matrix $A' = \sqrt{m/N}\, A^* \in \mathbb{R}^{N\times m}$ by changing m to N and s to m. Using the adjusted version of (24), we derive that, with probability at least $1 - 2\exp(-c_1 N)$,

$$\|A^*(\mathbf{e} - \mathbf{e}_i)\|_1 = \sqrt{\frac{N}{m}}\|A'(\mathbf{e} - \mathbf{e}_i)\|_1 \leq \sqrt{\frac{N}{m}}(1+\delta_*) \mathbin{/\!\!/} \mathbf{e} - \mathbf{e}_i \mathbin{/\!\!/} \leq \sqrt{\frac{N}{m}}\, 2\sqrt{N}\|\mathbf{e} - \mathbf{e}_i\|_2 \leq \frac{2N}{\sqrt{m}}\varepsilon.$$

In this situation, we obtain

$$d\sqrt{s_*}\|\!|A^*\mathbf{e}_i\|\!| \leq 1 + \frac{dN}{h}\sqrt{\frac{s_*}{m}}\,\varepsilon \leq 1 + \frac{dN}{h}\varepsilon = \left(1 + \frac{dN}{h}\varepsilon\right)\|\mathbf{e}_i\|_2.$$

Therefore, we have $d'\sqrt{s_*}\|\!|A^*\mathbf{e}_i\|\!| \leq \|\mathbf{e}_i\|_2$ once we have made the choice

$$\varepsilon := \frac{h}{dN}. \tag{30}$$

The previous considerations show that

$$\begin{aligned}
\Pr\big(\|\mathbf{e}\|_2 > d\sqrt{s_*}\|\!|A^*\mathbf{e}\|\!|,\ \text{some } \mathbf{e} \in \mathbb{R}^m\big) &\leq 2\exp(-c_1N) + \Pr\big(\|\mathbf{e}_i\|_2 > d'\sqrt{s_*}\|\!|A^*\mathbf{e}_i\|\!|,\ \text{some } i \in [k]\big)\\
&\leq 2\exp(-c_1N) + k\binom{2h}{h}\exp\left(-\frac{\sqrt{mN}}{3\sqrt{e}}\right) \leq 2\exp(-c_1N) + \left(\frac{3}{\varepsilon}\right)^m 4^h \exp\left(-\frac{\sqrt{mN}}{5}\right)\\
&= 2\exp(-c_1N) + \exp\left(m\ln\left(\frac{3}{\varepsilon}\right) + h\ln(4) - \frac{\sqrt{mN}}{5}\right)\\
&= 2\exp(-c_1N) + \exp\left(m\ln\left(\frac{3dN}{h}\right) + h\ln(4) - \frac{\sqrt{mN}}{5}\right).
\end{aligned}$$

Choosing $h = \sqrt{mN}/c$ with $c = 15/\ln(4)$ ensures that $h \leq N/4$. Since $N \geq C_2 m$, choosing the constant C_2 large enough to have $\ln(3dc\sqrt{C_2})/\sqrt{C_2} \leq 1/15$ also ensures that

$$\begin{aligned}
\exp\left(m\ln\left(\frac{3dN}{h}\right) + h\ln(4) - \frac{\sqrt{mN}}{5}\right) &= \exp\left(-\left[-\frac{\ln(3dc\sqrt{N/m})}{\sqrt{N/m}} - \frac{\ln(4)}{c} + \frac{1}{5}\right]\sqrt{mN}\right)\\
&\leq \exp\left(-\left[-\frac{\ln(3dc\sqrt{C_2})}{\sqrt{C_2}} - \frac{1}{15} + \frac{1}{5}\right]\sqrt{mN}\right) \leq \exp\left(-\frac{\sqrt{mN}}{15}\right).
\end{aligned}$$

We finally deduce that

$$\Pr\big(\|\mathbf{e}\|_2 > d\sqrt{s_*}\|\!|A^*\mathbf{e}\|\!|,\ \text{some } \mathbf{e} \in \mathbb{R}^m\big) \leq 2\exp(-c_1N) + \exp\left(-\frac{\sqrt{mN}}{15}\right) \leq 3\exp(-C_1\sqrt{mN}),$$

where $C_1 = \min\{c_1, 1/15\}$. $\square$

Acknowledgements: I thank Paweł Hitczenko for bringing the article [10] to my attention.

References

[1] R. ADAMCZAK, A. LITVAK, A. PAJOR, AND N. TOMCZAK–JAEGERMANN, *Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling*, Constructive Approximation, to appear.

[2] E. J. CANDÈS, Y. ELDAR, D. NEEDELL, AND P. RANDALL, *Compressed sensing with coherent and redundant dictionaries*, Appl. Comput. Harmon. Anal. 31 (2011) 59–73.

[3] E. J. CANDÈS, J. ROMBERG, AND T. TAO, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., 59 (2005) 1207–1223.

[4] A. COHEN, W. DAHMEN, AND R. DEVORE, *Compressed sensing and best k-term approximation*, JAMS 22 (2009) 211–231.

[5] R. DEVORE, G. PETROVA, AND P. WOJTASZCZYK, *Instance optimality in probability with an ℓ_1-minimalization decoder*, Appl. Comput. Harm. Analysis, 27 (2009) 275–288.

[6] S. FOUCART, *Sparse recovery algorithms: sufficient conditions in terms of restricted isometry constants*, Approximation Theory XIII: San Antonio 2010, Springer Proceedings in Mathematics.

[7] S. FOUCART AND M.-J. LAI, *Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$*, Appl. Comput. Harmon. Anal. 26 (2009) 395–407.

[8] S. FOUCART AND M.-J. LAI, *Sparse recovery with pre-Gaussian random matrices*, Studia Mathematica 200 (2010) 91–102.

[9] A. C. GILBERT, M. J. STRAUSS, J. A. TROPP, AND R. VERSHYNIN, *One sketch for all: fast algorithms for compressed sensing*, Proc. 39th ACM Symp. Theory of Computing.

[10] E. D. GLUSKIN AND S. KWAPIEŃ, *Tail and moment estimates for sums of independent random variables with logarithmically concave tails*, Studia Mathematica 114 (1995) 303–309.

[11] U. HAAGERUP, *The best constants in the Khinchine inequality*, Studia Mathematica 70 (1982) 231–283.

[12] G. G. LORENTZ, M. VON GOLITSCHEK, AND Y. MAKOVOZ, *Constructive Approximation: Advanced Problems*, Springer, Berlin, 1996.

[13] Q. MO AND S. LI, *New bounds on the restricted isometry constant δ_{2k}*, Appl. Comput. Harm. Analysis, 31 (2011) 460–468.

[14] A. E. LITVAK, A. PAJOR, M. RUDELSON, AND N. TOMCZAK–JAEGERMANN, *Smallest singular value of random matrices and geometry of random polytopes*, Advances in Mathematics, 195 (2005) 491–523.

[15] P. WOJTASZCZYK, *Stability and Instance Optimality for Gaussian Measurements in Compressed Sensing*, Found. Comput. Math. 10 (2010) 1–13.

[16] P. WOJTASZCZYK, ℓ_1 *minimization with noisy data*, SIAM J. Numer. Anal. 50 (2012) 458–467.