Supplementary Material for the ICML 2021 Publication:
A Nullspace Property for Subspace-Preserving Recovery

1 Introduction

A note on numbering. When we refer to a theorem/proposition/equation/etc. in this Supplementary Material, if its number has a section number in it, then it refers to a theorem/proposition/equation/etc. in the Supplementary Material. Otherwise it is a result in the original paper.

2 Preliminaries and problem formulation

Detailed notation and preliminaries. The set of integers from 1 up to $N$ is denoted as $[N] := \{1, \ldots, N\}$. For any $c \in \mathbb{R}^N$, the support of $c$ is denoted as $\text{Supp}(c) := \{k \in [N]: c_k \neq 0\}$. The vector $c$ is called $s$-sparse if $|\text{Supp}(c)| \leq s$. For any index set $S \subseteq [N]$, the complement of $S$ in $[N]$ is denoted by $S^c$. For a nonempty set $S \subseteq [N]$, the vector $c_S \in \mathbb{R}^{|S|}$ denotes the part of $c$ that is supported on $S$. We use $\text{Pr}_S \in \mathbb{R}^{N \times N}$ to denote the matrix that projects onto the coordinates in $S$ and sets all other coordinates to zero. For a matrix $X \in \mathbb{R}^{D \times N}$ and an index set $S \subseteq [N]$, the matrix $X_S \in \mathbb{R}^{D \times |S|}$ denotes the submatrix of $X$ consisting of the columns of $X$ indexed by $S$. Therefore, for all $c \in \mathbb{R}^N$, we have $X \text{Pr}_S c = X_S c_S$. If $S = \{j\}$ for some $j$, we simply write $x_j$ instead of $X_S$, to refer to the $j$th-column of $X$. We prioritize the subscript over superscript in the sense that $X_S^1 \equiv (X_S)^1$, and not $(X^1)_S$. Finally, $\text{Null}(X)$ denotes the nullspace of the matrix $X$ and $X^{-1}(\cdot)$ denotes the inverse image under $X$.

The $\ell_p$-norm of a vector $x \in \mathbb{R}^D$ is defined as $\|x\|_p := (\sum_{k=1}^D |x_k|^p)^{\frac{1}{p}}$, where $|\cdot|$ denotes the absolute value. The unit $\ell_p$-sphere is denoted by $B^D_p := \{x \in \mathbb{R}^D : \|x\|_p = 1\}$ and the unit $\ell_p$-ball is denoted by $B^D_p := \{x \in \mathbb{R}^D : \|x\|_p \leq 1\}$.

The convex hull is denoted by $\text{conv}(\cdot)$. We denote the convex hull of the union of the columns of $X$ and $-X$ by $K(X)$. Sometimes we refer to it as the symmetric convex hull of the columns of $X$. For a nonempty convex set $C \subseteq \mathbb{R}^D$, the set of extreme points of $C$ is denoted as $\text{Ext}(C)$. These are precisely the points that cannot be written as a nontrivial convex combination of two distinct points in $C$. The interior of $C$ is given by $\text{inte}(C) := \{x \in C : \exists \varepsilon > 0 \text{ s.t. } x + \varepsilon B^D_p \subseteq C\}$. Note that according to this definition, the interior of $C$ can be empty, although $C$ is non-empty. The affine hull of $C$, denoted by $\text{aff}(C)$, is the smallest affine set in $\mathbb{R}^D$ that contains $C$. The relative interior [1 p.44] of $C$ is defined as $\text{rinte}(C) := \{x \in \text{aff}(C) : \exists \varepsilon > 0, (x + \varepsilon B^D_p) \cap \text{aff}(C) \subseteq C\}$. The polar [1 p.125] of $C$ is defined as

$$C^\circ := \{q \in \mathbb{R}^D : q^\top x \leq 1 \text{ for all } x \in C\}. \quad (2.1)$$

Note $C^\circ$ is always a closed, convex set [1 p.125]. For any $\varepsilon > 0$, we have $(\varepsilon C)^\circ = \frac{1}{\varepsilon} C^\circ$ [1 Cor. 16.1.2]. If $C$ is closed and contains the origin, then $C^\circ$ is compact if and only if $0 \in \text{inte}(\text{conv}(C))$ [1 Cor. 14.5.1]. If $C_1, C_2 \subseteq \mathbb{R}^D$ are closed convex sets, then $C_1 \subseteq C_2$ if and only if $C_2 \subseteq C_1^\circ$ [1 p.125]. We define the dual $\|\cdot\|_*$ of a norm $\|\cdot\|$ as $\|z\|_* := \sup_{\|y\| \leq 1} z^\top y$. Then, if $\ell_q$ denotes the dual of $\ell_p$-norm, we have $\frac{1}{p} + \frac{1}{q} = 1$ and $B^D_q = (B^D_p)^\circ$.

We define the inner-$\ell_p$-radius of a nonempty compact convex set $C \subseteq \mathbb{R}^D$ containing the origin as the radius of the largest $\ell_p$-ball (confined to the linear span of $C$) one can inscribe inside $C$, and denote it by $r_p(C)$. That is, $r_p(C) := \max\{\alpha \in \mathbb{R}_{>0} : \alpha (B^D_p \cap \text{span}(C)) \subseteq C\}$, where $\text{span}(C)$ denotes the subspace spanned by $C$. Likewise, we define the outer-$\ell_p$-radius of $C$ as the radius of the smallest $\ell_p$-ball that contains $C$, and denote it by $R_p(C)$. That is, $R_p(C) := \min\{\beta \in \mathbb{R}_{>0} : \beta B^D_p \supseteq C\}$. 


3 A nullspace property for subspace-preserving recovery

Proof of Lemma [1]

Proof. Since each column of \( X_P \) can be written as a convex combination of the union of the columns of \( X_P \) and \(-X_P\), we conclude that \( X_P = X_P Y_P \) for some \( Y_P \in \mathbb{R}^{[P] \times |P|} \) with columns in \( S_1^{[P]-1} \). Also \( X_P = X_P \) holds trivially.

\[ \Rightarrow: \] Suppose that \( X \) satisfies SNSP. Let \( \tilde{\eta} \in \text{Null}(\tilde{X}, \tilde{P}) \) and \( \tilde{P} \in \tilde{P} \). There exists a unique \( P \in P \) such that \( \tilde{P} \subseteq P \). We lift \( \tilde{\eta} \) to \( \eta \in \text{Null}(X, P) \) by inserting zeros at the missing indices. Since \( X \) satisfies SNSP, the problem \( \min_{c \in X_P(\eta_P) = X_P(e)} \|c\|_1 \) has a minimizer \( \hat{c} \) which satisfies

\[ \|\hat{c}\|_1 < \|\eta_P\|_1 \tag{3.2} \]

Furthermore, we have

\[ \tilde{X}_P(Y_P \hat{c}) = (\tilde{X}_P Y_P) \hat{c} = X_P \hat{c} = X_P \eta_P = \tilde{X}_P \tilde{\eta}_P, \tag{3.3} \]

and we conclude that

\[ \min_{c \in X_P(\eta_P)} \|c\|_1 \leq \|Y_P \hat{c}\|_1 \leq \|\hat{c}\|_1 < \|\eta_P\|_1 \quad \text{By \eqref{3.2}} \]

and so, \( \tilde{X} \) satisfies SNSP.

\[ \Leftarrow: \] Conversely, suppose \( \tilde{X} \) satisfies SNSP. Let \( \eta \in \text{Null}(X, P) \). Suppose that \( X \) has \( r \) columns, and define \( \tilde{\eta} \in \mathbb{R}^r \) as the vector satisfying \( \tilde{\eta}_P := Y_P \eta_P \) for all \( P \in P \). Note that we have

\[ \tilde{X}_P(\tilde{\eta}_P) = (\tilde{X}_P Y_P) \eta_P = X_P \eta_P. \quad \text{\eqref{3.4}} \]

Hence, \( \tilde{X} \tilde{\eta} = \sum_{P \in \tilde{P}} \tilde{X}_P(\tilde{\eta}_P) = \sum_{P \in P} X_P \eta_P = X \eta = 0 \). So, we conclude that \( \tilde{\eta} \in \text{Null}(\tilde{X}) \).

If \( X_P \eta_P = 0 \) for all \( P \in P \), we have \( \min_{c \in X_P(\eta_P) = X_P(e)} \|c\|_1 = 0 < \|\eta_P\|_1 \) for all \( P \in P \). Therefore, \( \eta \) is satisfies for \( \eta \) and for all \( P \in \mathbb{P} \) trivially. So, w.l.o.g. we can assume that there exists \( Q, T \in \mathbb{P} \) with \( Q \neq T \), such that \( X_Q \eta_Q \neq 0 \neq X_T \eta_T \). In turn, we have \( \tilde{X}_Q \tilde{\eta}_Q \neq 0 \neq \tilde{X}_T \tilde{\eta}_T \) by \eqref{3.4}. Hence, w.l.o.g. \( \text{Supp}(\tilde{\eta}) \) is not contained in \( \tilde{P} \) for any \( \tilde{P} \in \tilde{P} \), and we have \( \tilde{\eta} \not\in \text{Null}(\tilde{X}, \tilde{P}) \).

Now we let \( P \in \mathbb{P} \), and argue as follows:

\[ \min_{c \in X_P(\eta_P) = X_P(e)} \|c\|_1 = \min_{c \in (X_P Y_P)(\eta_P) = X_P(e)} \|c\|_1 \leq \min_{z \in (\tilde{X}_P(\tilde{\eta}_P)) = (\tilde{X}_P(z))} \|z\|_1 \quad \text{By restricting the constraint set.} \]

\[ = \sum_{Q \neq P, Q \in \tilde{P}} \|\tilde{\eta}_Q\|_1 \]
\[ = \sum_{Q \neq P, Q \in \tilde{P}} \|Y_Q \eta_Q\|_1 \leq \sum_{Q \neq P, Q \in \tilde{P}} \|\eta_Q\|_1 \quad \text{Since } Y_Q \text{ has normalized columns.} \]
\[ = \|\eta_P\|_1. \]

Hence, \( X \) satisfies SNSP. \( \square \)

4 A geometrically interpretable characterization of SNSP

There are no extra proofs for this section, or any additional material of other kind.
5 Reduction of the verification of SNSP to a decision on finite sets

Reformulation of SNSP

Lemma 5.1. For any $P \in \mathcal{P}$, the function $f_P : \mathbb{R}^N \rightarrow \mathbb{R}_{\geq 0}$ defined as

$$f(\eta) = \|\eta P\|_1 + \min_{z : X_P z = X_P \eta P} \|z\|_1$$

is convex and positively homogeneous of degree 1.

Proof. We first aim to establish that $f_P$ is a convex function. Let $\lambda \in [0, 1]$ and $w, y \in \mathbb{R}^N$. Then,

$$f_P(\lambda w + (1 - \lambda)y) = \|\lambda w_P + (1 - \lambda)y_P\|_1 + \min_{z : X_P (\lambda w_P + (1 - \lambda)y_P) = X_P (z)} \|z\|_1$$

$$\leq \|\lambda w_P + (1 - \lambda)y_P\|_1 + \|\lambda(y_P + \eta_w) + (1 - \lambda)(y_P + \eta_y)\|_1 \quad \text{for all } \eta_w, \eta_y \in \text{Null}(X_P)$$

$$\leq \lambda(\|w_P\|_1 + \|w_P + \eta_w\|_1) + (1 - \lambda)(\|y_P\|_1 + \|y_P + \eta_y\|_1) \quad \text{for all } \eta_w, \eta_y \in \text{Null}(X_P).$$

In particular,

$$f_P(\lambda w + (1 - \lambda)y) \leq \lambda\left(\|w_P\|_1 + \min_{z : X_P (\lambda w_P) = X_P (z)} \|z\|_1\right) + (1 - \lambda)\left(\|y_P\|_1 + \min_{z : X_P (y_P) = X_P (z)} \|z\|_1\right)$$

$$= \lambda f_P(w) + (1 - \lambda)f_P(y),$$

which establishes that $f_P$ is a convex function.

In order to show positive homogeneity, let $\alpha \in \mathbb{R}_{> 0}$. Then, it follows that

$$f_P(\alpha w) = \|\alpha w_P\|_1 + \min_{z : X_P (\alpha w_P) = X_P (z)} \|z\|_1$$

$$= \alpha \|w_P\|_1 + \min_{z : X_P (\alpha w_P) = X_P (z)} \|z\|_1$$

$$= \alpha \|w_P\|_1 + \alpha \min_{y : X_P (\alpha w_P) = X_P (y)} \|y\|_1$$

$$= \alpha \|w_P\|_1 + \alpha \min_{y : X_P (w_P) = X_P (y)} \|y\|_1 = \alpha f_P(w),$$

which completes the proof. $\square$

Proposition 5.1. The matrix $X$ satisfies SNSP if and only if for all $\eta \in \text{Null}(X, \mathcal{P}) \cap \mathbb{B}_1^N$ and for all $P \in \mathcal{P}$, we have

$$\|\eta P\|_1 + \min_{z : X_P (\eta P) = X_P (z)} \|z\|_1 < 1 \quad (5.6)$$

Proof. The matrix $X$ satisfies SNSP if and only if $f_P(\eta) < \|\eta\|_1$. That is, if and only if, $f_P\left(\frac{\eta}{\|\eta\|_1}\right) < 1$, since $f_P$ is positively homogeneous by Lemma 5.1. Hence, $X$ satisfies SNSP if and only if for all $\eta \in \text{Null}(X, \mathcal{P}) \cap \mathbb{B}_1^N$ and for all $P \in \mathcal{P}$, we have $f_P(\eta) < 1$. This is true, if and only, for all $\eta \in \text{Null}(X, \mathcal{P}) \cap \mathbb{B}_1^N$ and for all $P \in \mathcal{P}$, (5.6) holds, as claimed. $\square$

Proof of Thm. 3

Proof. We will use Prop. 5.1 to show the equivalence. For this purpose, what we need to show is the following: The inequality (5.6) holds for all $P \in \mathcal{P}$ and for all $\eta \in \text{Null}(X, \mathcal{P}) \cap \mathbb{B}_1^N$, if and only if, it holds for all $P \in \mathcal{P}$ and for all $\eta \in \text{Ext}(\text{Null}(X) \cap \mathbb{B}_1^N, \mathcal{P})$.

$\Rightarrow$: Obvious, since $\text{Ext}(\text{Null}(X) \cap \mathbb{B}_1^N, \mathcal{P}) \subset \text{Null}(X, \mathcal{P}) \cap \mathbb{B}_1^N$. 3
\( \iff \) Let \( P \in \mathcal{P} \), and \( \eta \in \text{Null}(X, \mathcal{P}) \cap B^N_1 \). Then, there exists \( r \geq 1 \), \( \{w_i\}_{i=1}^r \in \text{Ext}(\text{Null}(X) \cap B^N_1) \) and \( \{\lambda_i\}_{i=1}^r \subset (0, 1] \) with \( \sum_{i=1}^r \lambda_i = 1 \) such that \( \eta = \sum_{i=1}^r \lambda_i w_i \).

Let \( f_P \) be the function in Lemma 5.1. Note that for all \( l \in \{1, \ldots, r\} \), if there exists \( Q_l \in \mathcal{P} \) such that \( \text{Supp}(w_l) \subseteq Q_l \), then

\[
 f_P(w_l) = \begin{cases} 
 0 & \text{if } Q_l \neq P, \\
 1 & \text{if } Q_l = P.
\end{cases}
\]

If no such \( Q_l \) exists, then \( w_l \in \text{Ext}(\text{Null}(X) \cap B^N_1, \mathcal{P}) \), and \( f_P(w_l) < 1 \), by our hypothesis. So, we conclude that \( f_P(w_l) < 1 \), if \( \text{Supp}(w_l) \not\subseteq P \).

Note that there must exist \( l \in \{1, \ldots, r\} \) such that \( \text{Supp}(w_l) \not\subseteq P \) because, otherwise, \( \text{Supp}(\eta) \subseteq P \), which would be a contradiction. Moreover, \( f_P(w_l) < 1 \), which we can use to argue that

\[
 \|\eta\|_1 + \min_{z : X_P(\eta) = X_P(z)} \|z\|_1 = f_P(\eta) = f_P\left( \sum_{i=1}^r \lambda_i w_i \right) 
\leq \sum_{i=1}^r \lambda_i f_P(w_l) \quad \text{Since } f_P \text{ is convex by Lemma 5.1}
\leq \sum_{i=1}^r \lambda_i = 1 \quad \text{Since } f_P(w_l) < 1 \text{ and } f_P(w_l) \leq 1,
\]

which completes the proof. \( \square \)

**Auxiliary results for the dual of Basis Pursuit**

When \( \Psi \in \mathbb{R}^{r \times s} \), the dual of the \( \ell_1 \)-minimization problem

\[
 \min_{\Psi \in \mathbb{R}^{r \times s}} \|y\|_1
\]

is given by the following two equivalent forms

\[
 \max_{\Psi \in \mathbb{R}^{r \times s}} \bar{y}^T \Psi v \equiv \max_{w \in \text{im}(\Psi^\top) \cap \mathbb{B}_s^\infty} \bar{y}^T w.
\]

\[
 (5.7)
\]

Since there is no duality gap and we have \( \min_{\Psi \in \mathbb{R}^{r \times s}} \|y\|_1 = \max_{w \in \text{im}(\Psi^\top) \cap \mathbb{B}_s^\infty} \bar{y}^T w \). We use the dual problem to derive sufficient conditions for SNSP, which are geometrically more interpretable. Therefore, the structure of \( \text{im}(\Psi^\top) \cap \mathbb{B}_s^\infty \) is of particular interest, and the next result helps us understand it better, when the columns of \( \Psi \) have unit \( \ell_p \)-norm.

**Lemma 5.2.** Let \( \Psi \in \mathbb{R}^{r \times s} \) be a matrix with columns of unit \( \ell_p \)-norm with \( p \in [1, \infty] \), and \( \ell_q \) be the dual of \( \ell_p \). Then,

\[
 \mathbb{B}_1^s \subseteq \text{im}(\Psi) \cap \mathbb{B}_p^r,
\]

\[
 (5.9)
\]

or alternatively,

\[
 \text{im}(\Psi^\top) \cap \mathbb{B}_s^\infty \supseteq \Psi^\top \mathbb{B}_q^r.
\]

\[
 (5.10)
\]

When \( p \in \{1, \infty\} \), the equality holds in \((5.9)\), if and only if, equality holds in \((5.10)\).

**Proof.** Since the columns of \( \Psi \) have unit \( \ell_p \)-norm, we must have \( \Psi v \in \mathbb{B}_p^r \) for all \( v \in \mathbb{B}_1^s \). That is, we have \( \Psi^\top \mathbb{B}_1^s \subseteq \mathbb{B}_p^r \), so that \((5.9)\) follows trivially.

The inclusion \((5.9)\) implies \( \mathbb{B}_1^s \subseteq \Psi^{-1} \mathbb{B}_p^r \). Taking the polar of both sides, we obtain \( \mathbb{B}_s^\infty = (\mathbb{B}_1^s)^0 \supseteq (\Psi^{-1} \mathbb{B}_p^r)^0 = \Psi^T \mathbb{B}_q^r \) by \([\text{H] Cor. } 16.3.2]\). Since \( \text{im}(\Psi^\top) \supseteq \Psi^T \mathbb{B}_q^r \) holds trivially, we obtain \((5.10)\).

Suppose that \( p \in \{1, \infty\} \) and equality holds in \((5.9)\). Taking inverse image under \( \Psi \) on both sides of \((5.9)\), we obtain \( \Psi^{-1} \Psi \mathbb{B}_1^s = \Psi^{-1} (\text{im}(\Psi) \cap \mathbb{B}_p^r) \). The left-hand-side of this equation is \( \text{Null}(\Psi) + \mathbb{B}_1^s \), whereas the right-hand-side is equal to \( \Psi^{-1} \mathbb{B}_p^r \). That is, we have \( \text{Null}(\Psi) + \mathbb{B}_1^s = \Psi^{-1} \mathbb{B}_p^r \). Now, taking polar of both
expressive, if and only if, one of the following equivalent conditions hold:

1. We set

\[ \text{Proof.} \]

Lemma 5.3. Let \( \Psi \in \mathbb{R}^{n \times s} \) be a matrix with columns of unit \( \ell_p \)-norm. We say that \( \Psi \) is perfectly expressive, if \( r_p(\Psi B_1^s) = 1 \).

We also have the following lemma, which characterizes perfectly expressiveness.

Lemma 5.3. Let \( \Psi \in \mathbb{R}^{n \times s} \) be a matrix with columns of unit \( \ell_p \)-norm with \( p \in \{1, \infty\} \). Then, \( \Psi \) is perfectly expressive, if and only if, one of the following equivalent conditions hold:

1. \( \Psi B_1^s = \text{im}(\Psi) \cap \mathbb{B}_p^r \).

2. For all \( v \in \text{Ext}(\text{im}(\Psi) \cap \mathbb{B}_p^r) \), \( v \) or \( -v \) is a column of \( \Psi \).

3. \( \text{im}(\Psi^T) \cap \mathbb{B}_\infty^s = \Psi^T B_q^r \).

Proof. 1 \( \Rightarrow \) 2: If \( \Psi B_1^s = \text{im}(\Psi) \cap \mathbb{B}_p^r \), then we necessarily have \( \text{Ext}(\text{im}(\Psi) \cap \mathbb{B}_p^r) = \text{Ext}(\text{im}(\Psi) \cap \mathbb{B}_p^r) \). That is, any \( v \in \text{Ext}(\text{im}(\Psi) \cap \mathbb{B}_p^r) \) is an extreme point of \( \Psi B_1^s \). But each extreme point of \( \Psi B_1^s \) is either a column of \( \Psi \) or a column of \( -\Psi \), and we get the result.

2 \( \Rightarrow \) 1: If for all \( v \in \text{Ext}(\text{im}(\Psi) \cap \mathbb{B}_p^r) \), \( v \) or \(-v\) is a column of \( \Psi \), then \( \text{Ext}(\text{im}(\Psi) \cap \mathbb{B}_p^r) \subseteq \Psi B_1^s \). Taking convex hull, we obtain \( \text{im}(\Psi) \cap \mathbb{B}_p^r = \text{conv}(\text{Ext}(\text{im}(\Psi) \cap \mathbb{B}_p^r)) \subseteq \Psi B_1^s \). Since the converse inclusion always holds by Lemma 5.2, we obtain the equality.

3 \( \iff \) 1: Follows from Lemma 5.2. \( \square \)

We end this part with a final auxiliary result that will be useful in the proof of Prop. 1

Lemma 5.4. Let \( \Psi \in \mathbb{R}^{n \times s} \) be a matrix with columns of unit \( \ell_p \)-norm with \( p \in [1, \infty) \), and \( \ell_q \) be the dual of \( \ell_p \). Then,

\[
\frac{1}{r_p(\Psi B_1^s)} = \min\{\beta \in \mathbb{R}_{\geq 1} : \text{im}(\Psi^T) \cap \mathbb{B}_\infty^s \subseteq \beta \Psi^T B_q^r\}. \tag{5.11}
\]

Proof. We set \( \beta^* := \min\{\beta \in \mathbb{R}_{\geq 1} : \text{im}(\Psi^T) \cap \mathbb{B}_\infty^s \subseteq \beta \Psi^T B_q^r\} \). Then

\[
\text{im}(\Psi^T) \cap \mathbb{B}_\infty^s \subseteq \beta^* \Psi^T B_q^r \iff (\text{im}(\Psi^T) \cap \mathbb{B}_\infty^s) \subseteq (\beta^* \Psi^T B_q^r)^o \iff \text{Null}(\Psi) + \mathbb{E}_1^s \supseteq \frac{1}{\beta^*} \Psi^T B_q^r \iff \text{Null}(\Psi) + \beta^* \mathbb{E}_1^s \supseteq \Psi^T B_q^r \iff \beta^* \Psi B_1^s \supseteq \Psi^T B_q^r = \text{im}(\Psi^T) \cap \mathbb{B}_p^r.
\]

Hence, we must have \( \beta^* \geq \frac{1}{r_p(\Psi B_1^s)} \). Conversely, we have

\[
\text{im}(\Psi) \cap \mathbb{B}_p^r \subseteq \frac{1}{r_p(\Psi B_1^s)} \Psi B_1^s \iff (\text{im}(\Psi) \cap \mathbb{B}_p^r) \supseteq \left(\frac{1}{r_p(\Psi B_1^s)} \Psi B_1^s\right)^o \iff \text{Null}(\Psi^T) + B_q^r \supseteq r_p(\Psi B_1^s) (\Psi^T)^{-1} \mathbb{B}_q^s \iff \text{Null}(\Psi^T) + \frac{1}{r_p(\Psi B_1^s)} \Psi^T \mathbb{B}_q^r \supseteq (\Psi^T)^{-1} \mathbb{B}_q^s \iff \frac{1}{r_p(\Psi B_1^s)} \Psi^T B_q^r \supseteq (\Psi^T)^{-1} \mathbb{B}_q^s = \text{im}(\Psi^T) \cap \mathbb{B}_p^s.
\]

So, we also have the reverse inequality, \( \beta^* \leq \frac{1}{r_p(\Psi B_1^s)} \). Hence, \( \beta^* = \frac{1}{r_p(\Psi B_1^s)} \). \( \square \)
Proof of Prop. 1.

Proof. By Lemma 5.4, we have \( \text{im}(\Psi^\top) \cap B_\infty^s \subseteq \frac{1}{\tau_p(\Psi B_1^s)} \Psi^\top B_q^r \). Hence,

\[
\max_{w \in \text{im}(\Psi^\top) \cap B_\infty^s} \bar{y}^\top w \leq \max_{w \in \frac{1}{\tau_p(\Psi B_1^s)} \Psi^\top B_q^r} \bar{y}^\top w
\]

(5.12)

\[
= \frac{1}{\tau_p(\Psi B_1^s)} \max_{w \in \Psi^\top B_q^r} \bar{y}^\top w
\]

\[
= \frac{1}{\tau_p(\Psi B_1^s)} \max_{e \in B_q^r} \bar{y}^\top \Psi^\top e
\]

\[
= \frac{\|\Psi \bar{y}\|_p}{\tau_p(\Psi B_1^s)}.
\]

If \( p \in \{1, \infty\} \) and \( \tau_p(\Psi B_1^s) = 1 \) (i.e. \( \Psi \) is perfectly expressive), then, by Lemma 5.3, equality holds in (5.12) with \( \tau_p(\Psi B_1^s) = 1 \), and the result follows.

References