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

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

The convex hull is denoted by $\mathrm{conv}(\cdot)$. We denote the convex hull of the union of the columns of $\boldsymbol{X}$ and $-\boldsymbol{X}$ by $\mathcal{K}(\boldsymbol{X})$. Sometimes we refer to it as the *symmetrized convex hull* of the columns of $\boldsymbol{X}$. For a nonempty convex set $\mathcal{C} \subseteq \mathbb{R}^D$, the set of extreme points of $\mathcal{C}$ is denoted as $\mathrm{Ext}(\mathcal{C})$. These are precisely the points that cannot be written as a nontrivial convex combination of two distinct points in $\mathcal{C}$. The *interior* of $\mathcal{C}$ is given by $\mathrm{inte}(\mathcal{C}) := \{\boldsymbol{x} \in \mathcal{C} : \exists \varepsilon > 0 \text{ s.t. } \boldsymbol{x} + \varepsilon \mathbb{B}_1^D \subseteq \mathcal{C}\}$. Note that according to this definition, the interior of $\mathcal{C}$ can be empty, although $\mathcal{C}$ is non-empty. The *affine hull* of $\mathcal{C}$, denoted by $\mathrm{aff}(\mathcal{C})$, is the smallest affine set in $\mathbb{R}^D$ that contains $\mathcal{C}$. The *relative interior* [1, p.44] of $\mathcal{C}$ is defined as $\mathrm{rinte}(\mathcal{C}) := \{\boldsymbol{x} \in \mathrm{aff}(\mathcal{C}) : \exists \varepsilon > 0, (\boldsymbol{x} + \varepsilon \mathbb{B}_2^D) \cap \mathrm{aff}(\mathcal{C}) \subseteq \mathcal{C}\}$. The *polar* [1, p.125] of $\mathcal{C}$ is defined as

$$\mathcal{C}^\circ := \{\boldsymbol{q} \in \mathbb{R}^D : \boldsymbol{q}^\top \boldsymbol{x} \leq 1 \text{ for all } \boldsymbol{x} \in \mathcal{C}\}. \tag{2.1}$$

Note $\mathcal{C}^\circ$ is always a closed, convex set [1, p.125]. For any $\varepsilon > 0$, we have $(\varepsilon \mathcal{C})^\circ = \frac{1}{\varepsilon} \mathcal{C}^\circ$ [1, Cor. 16.1.2]. If $\mathcal{C}$ is closed and contains the origin, then $\mathcal{C}^\circ$ is compact if and only if $0 \in \mathrm{inte}(\mathrm{conv}(\mathcal{C}))$ [1, Cor. 14.5.1]. If $\mathcal{C}_1, \mathcal{C}_2 \subset \mathbb{R}^D$ are closed convex sets, then $\mathcal{C}_1 \subseteq \mathcal{C}_2$ if and only if $\mathcal{C}_2^\circ \subseteq \mathcal{C}_1^\circ$ [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 $\mathbb{B}_q^D = (\mathbb{B}_p^D)^\circ$.

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

# 3 A nullspace property for subspace-preserving recovery

## Proof of Lemma 1

*Proof.* Since each column of $\boldsymbol{X}_P$ can be written as a convex combination of the union of the columns of $\boldsymbol{X}_{\tilde{P}}$ and $-\boldsymbol{X}_{\tilde{P}}$, we conclude that $\boldsymbol{X}_P = \boldsymbol{X}_{\tilde{P}} Y_P$ for some $Y_P \in \mathbb{R}^{|\tilde{P}|\times|P|}$ with columns in $\mathbb{S}_1^{|\tilde{P}|-1}$. Also $\boldsymbol{X}_{\tilde{P}} = \tilde{\boldsymbol{X}}_{\tilde{P}}$ holds trivially.

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

$$\|\hat{c}\|_1 < \|\eta_{P^c}\|_1 \tag{3.2}$$

Furthermore, we have

$$\tilde{\boldsymbol{X}}_{\tilde{P}}(Y_P\hat{c}) = (\tilde{\boldsymbol{X}}_{\tilde{P}}Y_P)\hat{c} = \boldsymbol{X}_P\hat{c} = \boldsymbol{X}_P\eta_P = \tilde{\boldsymbol{X}}_{\tilde{P}}\tilde{\eta}_{\tilde{P}}, \tag{3.3}$$

and we conclude that

$$\begin{aligned}
\min_{\boldsymbol{c}:\tilde{\boldsymbol{X}}_{\tilde{P}}\tilde{\eta}_{\tilde{P}}=\tilde{\boldsymbol{X}}_{\tilde{P}}(\boldsymbol{c})} \|\boldsymbol{c}\|_1 &\le \|Y_P\hat{c}\|_1 && \text{Since } Y_P\hat{c} \text{ is feasible by (3.3).}\\
&\le \|\hat{c}\|_1 && \text{Since } Y_P \text{ has normalized columns.}\\
&< \|\eta_{P^c}\|_1 && \text{By (3.2)}\\
&= \|\tilde{\eta}_{\tilde{P}^c}\|_1,
\end{aligned}$$

and so, $\tilde{\boldsymbol{X}}$ satisfies SNSP.

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

$$\tilde{\boldsymbol{X}}_{\tilde{P}}(\tilde{\eta}_{\tilde{P}}) = (\tilde{\boldsymbol{X}}_{\tilde{P}}Y_P)\eta_P = \boldsymbol{X}_P\eta_P. \tag{3.4}$$

Hence, $\tilde{\boldsymbol{X}}\tilde{\eta} = \sum_{\tilde{P}\in\tilde{\mathcal{P}}} \tilde{\boldsymbol{X}}_{\tilde{P}}(\tilde{\eta}_{\tilde{P}}) = \sum_{P\in\mathcal{P}} \boldsymbol{X}_P\eta_P = \boldsymbol{X}\eta = 0$. So, we conclude that $\tilde{\eta} \in \mathrm{Null}(\tilde{\boldsymbol{X}})$.

If $\boldsymbol{X}_P\eta_P = 0$ for all $P \in \mathcal{P}$, we have $\min_{\boldsymbol{c}:\boldsymbol{X}_P(\eta_P)=\boldsymbol{X}_P(\boldsymbol{c})} \|\boldsymbol{c}\|_1 = 0 < \|\eta_{P^c}\|_1$ for all $P \in \mathcal{P}$. Therefore, (4) is satisfies for $\eta$ and for all $P \in \mathcal{P}$ trivially. So, w.l.o.g. we can assume that there exists $Q, T \in \mathcal{P}$ with $Q \neq T$, such that $\boldsymbol{X}_Q\eta_Q \neq 0 \neq \boldsymbol{X}_T\eta_T$. In turn, we have $\tilde{\boldsymbol{X}}_{\tilde{Q}}\tilde{\eta}_{\tilde{Q}} \neq 0 \neq \tilde{\boldsymbol{X}}_{\tilde{T}}\tilde{\eta}_{\tilde{T}}$ by (3.4). Hence, w.l.o.g. $\mathrm{Supp}(\tilde{\eta})$ is not contained in $\tilde{P}$ for any $\tilde{P} \in \tilde{\mathcal{P}}$, and we have $\tilde{\eta} \in \mathrm{Null}(\tilde{\boldsymbol{X}}, \tilde{\mathcal{P}})$.

Now we let $P \in \mathcal{P}$, and argue as follows:

$$\begin{aligned}
\min_{\boldsymbol{c}:\boldsymbol{X}_P(\eta_P)=\boldsymbol{X}_P(\boldsymbol{c})} \|\boldsymbol{c}\|_1 &= \min_{\boldsymbol{c}:(\tilde{\boldsymbol{X}}_{\tilde{P}}Y_P)(\eta_P)=\boldsymbol{X}_P(\boldsymbol{c})} \|\boldsymbol{c}\|_1\\
&\le \min_{\boldsymbol{z}:\tilde{\boldsymbol{X}}_{\tilde{P}}(\tilde{\eta}_{\tilde{P}})=\tilde{\boldsymbol{X}}_{\tilde{P}}(\boldsymbol{z})} \|\boldsymbol{z}\|_1 && \text{By restricting the constraint set.}\\
&< \|\tilde{\eta}_{\tilde{P}^c}\|_1 && \text{Since } \tilde{\boldsymbol{X}} \text{ satisfies SNSP.}\\
&= \sum_{\tilde{Q}\neq\tilde{P},\tilde{Q}\in\tilde{\mathcal{P}}} \|\tilde{\eta}_{\tilde{Q}}\|_1\\
&= \sum_{Q\neq P,Q\in\mathcal{P}} \|Y_Q\eta_Q\|_1\\
&\le \sum_{Q\neq P,Q\in\mathcal{P}} \|\eta_Q\|_1 && \text{Since } Y_Q \text{ has normalized columns.}\\
&= \|\eta_{P^c}\|_1.
\end{aligned}$$

Hence, $\boldsymbol{X}$ satisfies SNSP. □

# 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 $f_P : \mathbb{R}^N \to \mathbb{R}_{\geq 0}$ defined as*

$$f(\eta) = \|\eta_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P\boldsymbol{z}=\boldsymbol{X}_P\eta_P} \|\boldsymbol{z}\|_1 \tag{5.5}$$

*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,

$$\begin{aligned}
&f_P(\lambda w + (1-\lambda)y) \\
&\quad = \|\lambda w_P + (1-\lambda)y_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P(\lambda w_P+(1-\lambda)y_P)=\boldsymbol{X}_P(\boldsymbol{z})} \|\boldsymbol{z}\|_1 \\
&\quad \leq \|\lambda w_P + (1-\lambda)y_P\|_1 + \|\lambda(w_P+\eta_w) + (1-\lambda)(y_P+\eta_y)\|_1 \quad \text{for all } \eta_w, \eta_y \in \mathrm{Null}(\boldsymbol{X}_P) \\
&\quad \leq \lambda(\|w_P\|_1 + \|w_P+\eta_w\|_1) + (1-\lambda)(\|y_P\|_1 + \|y_P+\eta_y\|_1) \quad \text{for all } \eta_w, \eta_y \in \mathrm{Null}(\boldsymbol{X}_P).
\end{aligned}$$

In particular,

$$\begin{aligned}
&f_P(\lambda w + (1-\lambda)y) \\
&\quad \leq \lambda\left(\|w_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P(w_P)=\boldsymbol{X}_P(\boldsymbol{z})} \|\boldsymbol{z}\|_1\right) + (1-\lambda)\left(\|y_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P(y_P)=\boldsymbol{X}_P(\boldsymbol{z})} \|\boldsymbol{z}\|_1\right) \\
&\quad = \lambda f_P(w) + (1-\lambda) f_P(y),
\end{aligned}$$

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

$$\begin{aligned}
f_P(\alpha w) &= \|\alpha w_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P(\alpha w_P)=\boldsymbol{X}_P(\boldsymbol{z})} \|\boldsymbol{z}\|_1 \\
&= \alpha\|w_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P(w_P)=\boldsymbol{X}_P\left(\frac{1}{\alpha}\boldsymbol{z}\right)} \|\boldsymbol{z}\|_1 \\
&= \alpha\|w_P\|_1 + \min_{\boldsymbol{y}:\boldsymbol{X}_P(w_P)=\boldsymbol{X}_P(\boldsymbol{y})} \|\alpha\boldsymbol{y}\|_1 \\
&= \alpha\|w_P\|_1 + \alpha \min_{\boldsymbol{y}:\boldsymbol{X}_P(w_P)=\boldsymbol{X}_P(\boldsymbol{y})} \|\boldsymbol{y}\|_1 = \alpha f_P(w),
\end{aligned}$$

which completes the proof. □

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

$$\|\eta_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P(\eta_P)=\boldsymbol{X}_P(\boldsymbol{z})} \|\boldsymbol{z}\|_1 < 1 \tag{5.6}$$

*Proof.* The matrix $\boldsymbol{X}$ satisfies SNSP if and only if $\eta \in \mathrm{Null}(\boldsymbol{X}, \mathcal{P})$ and for all $P \in \mathcal{P}$, we have $f_P(\eta) < \|\eta\|_1$. 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, $\boldsymbol{X}$ satisfies SNSP if and only if for all $\eta \in \mathrm{Null}(\boldsymbol{X}, \mathcal{P}) \cap \mathbb{S}_1^{N-1}$ and for all $P \in \mathcal{P}$, we have $f_P(\eta) < 1$. This is true, if and only, for all $\eta \in \mathrm{Null}(\boldsymbol{X}, \mathcal{P}) \cap \mathbb{B}_1^N$ and for all $P \in \mathcal{P}$ (5.6) holds, as claimed. □

### 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 $\eta \in \mathrm{Null}(\boldsymbol{X}, \mathcal{P}) \cap \mathbb{B}_1^N$, if and only if, it holds for all $P \in \mathcal{P}$ and for all $\eta \in \mathrm{Ext}(\mathrm{Null}(\boldsymbol{X}) \cap \mathbb{B}_1^N, \mathcal{P})$.

$\Rightarrow$: Obvious, since $\mathrm{Ext}(\mathrm{Null}(\boldsymbol{X}) \cap \mathbb{B}_1^N, \mathcal{P}) \subset \mathrm{Null}(\boldsymbol{X}, \mathcal{P}) \cap \mathbb{B}_1^N$.

$\Leftarrow$: Let $P \in \mathcal{P}$, and $\eta \in \text{Null}(\boldsymbol{X}, \mathcal{P}) \cap \mathbb{B}_1^N$. Then, there exists $r \geq 1$, $\{w_l\}_{l=1}^r \in \text{Ext}(\text{Null}(\boldsymbol{X}) \cap \mathbb{B}_1^N)$ and $\{\lambda_l\}_{l=1}^r \subset (0, 1]$ with $\sum_{l=1}^r \lambda_l = 1$ such that $\eta = \sum_{l=1}^r \lambda_l w_l$.

Let $f_P$ be the function in Lemma 5.1. Note that for all $l \in \{1, \cdots, r\}$, 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 $w_l \in \text{Ext}(\text{Null}(\boldsymbol{X}) \cap \mathbb{B}_1^N, \mathcal{P})$, 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 $\bar{l} \in \{1, \ldots, r\}$ such that $\text{Supp}(w_{\bar{l}}) \not\subseteq P$ because, otherwise, $\text{Supp}(\eta) \subseteq P$, which would be a contradiction. Moreover, $f_P(w_{\bar{l}}) < 1$, which we can use to argue that

$$\begin{aligned}
&\|\eta_P\|_1 + \min_{\boldsymbol{z}:\boldsymbol{X}_P(\eta_P) = \boldsymbol{X}_P(\boldsymbol{z})} \|\boldsymbol{z}\|_1 \\
&\quad = f_P(\eta) = f_P\left(\sum_{l=1}^r \lambda_l w_l\right) \\
&\quad \leq \sum_{l=1}^r \lambda_l f_P(w_l) && \text{Since } f_P \text{ is convex by Lemma 5.1} \\
&\quad < \sum_{l=1}^r \lambda_l = 1 && \text{Since } f_P(w_{\bar{l}}) < 1 \text{ and } f_P(w_l) \leq 1,
\end{aligned}$$

which completes the proof. □

## Auxiliary results for the dual of Basis Pursuit

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

$$\min_{\Psi \bar{y} = \Psi y} \|y\|_1 \tag{5.7}$$

is given by the following two equivalent forms

$$\max_{\Psi^\top v \in \mathbb{B}_\infty^s} \bar{y}^\top \Psi^\top v \qquad \equiv \qquad \max_{w \in \text{im}(\Psi^\top) \cap \mathbb{B}_\infty^s} \bar{y}^\top w. \tag{5.8}$$

Since there is no duality gap and we have $\min_{\Psi \bar{y} = \Psi y} \|y\|_1 = \max_{w \in \text{im}(\Psi^\top) \cap \mathbb{B}_\infty^s} \bar{y}^\top w$. We use the dual problem to derive sufficient conditions for SNSP, which are geometrically more interpretable. Therefore, the structure of $\text{im}(\Psi^\top) \cap \mathbb{B}_\infty^s$ 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,*

$$\Psi \mathbb{B}_1^s \subseteq \text{im}(\Psi) \cap \mathbb{B}_p^r, \tag{5.9}$$

*or alternatively,*

$$\text{im}(\Psi^\top) \cap \mathbb{B}_\infty^s \supseteq \Psi^\top \mathbb{B}_q^r. \tag{5.10}$$

*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 $\Psi \mathbb{B}_1^s \subseteq \mathbb{B}_p^r$, 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 $\mathbb{B}_\infty^s = (\mathbb{B}_1^s)^\circ \supseteq \left(\Psi^{-1} \mathbb{B}_p^r\right)^\circ = \Psi^T \mathbb{B}_q^r$ by [1, 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 $\Psi^{-1} \Psi \mathbb{B}_1^s = \Psi^{-1} \left(\text{im}(\Psi) \cap \mathbb{B}_p^r\right)$. 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 sides, we obtain $\mathrm{im}(\Psi^\top) \cap \mathbb{B}_\infty^s = (\mathrm{Null}(\Psi) + \mathbb{B}_1^s)^\circ = \left(\Psi^{-1}\mathbb{B}_p^r\right)^\circ = \Psi^\top \mathbb{B}_q^r$, by [1, Cor. 16.3.2 & Cor. 16.5.2]. Conversely, suppose that equality holds in (5.10). Then, by taking the polar of both sides of the equality, we obtain $\mathrm{Null}(\Psi) + \mathbb{B}_1^s = \Psi^{-1}\mathbb{B}_p^r$, by [1, Cor. 9.1.1, Cor. 16.3.2 & Cor. 16.5.2]. Applying $\Psi$ to both sides, we obtain $\Psi\mathbb{B}_1^s = \mathrm{im}(\Psi) \cap \mathbb{B}_p^r$. □

Since the span of $\Psi\mathbb{B}_1^s$ is $\mathrm{im}(\Psi)$, the definition of inner-$\ell_p$-radius implies that $\mathfrak{r}_p\left(\Psi\mathbb{B}_s^1\right)$ is the largest number such that $\frac{1}{\mathfrak{r}_p(\Psi\mathbb{B}_s^1)}\Psi\mathbb{B}_1^s \supseteq \mathrm{im}(\Psi) \cap \mathbb{B}_p^r$. When $p \notin \{1, \infty\}$, we always have $\mathfrak{r}_p\left(\Psi\mathbb{B}_s^1\right) < 1$, but when $p \in \{1, \infty\}$, the equality can be attained, so we give this case a special name.

**Definition 5.1.** *Let $p \in \{1, \infty\}$ and $\Psi \in \mathbb{R}^{r\times s}$ be a matrix with columns of unit $\ell_p$-norm. We say that $\Psi$ is* perfectly expressive, *if $\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right) = 1$.*

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

**Lemma 5.3.** *Let $\Psi \in \mathbb{R}^{r\times s}$ 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. $\Psi\mathbb{B}_1^s = \mathrm{im}(\Psi) \cap \mathbb{B}_p^r$.
2. *For all $v \in \mathrm{Ext}(\mathrm{im}(\Psi) \cap \mathbb{B}_p^r)$, $v$ or $-v$ is a column of $\Psi$.*
3. $\mathrm{im}(\Psi^\top) \cap \mathbb{B}_\infty^s = \Psi^\top\mathbb{B}_q^r$.

*Proof.* $1 \Rightarrow 2$ : If $\Psi\mathbb{B}_1^s = \mathrm{im}(\Psi) \cap \mathbb{B}_p^r$, then we necessarily have $\mathrm{Ext}(\Psi\mathbb{B}_1^s) = \mathrm{Ext}(\mathrm{im}(\Psi) \cap \mathbb{B}_p^r)$. That is, any $v \in \mathrm{Ext}(\mathrm{im}(\Psi) \cap \mathbb{B}_p^r)$ is an extreme point of $\Psi\mathbb{B}_1^s$. But each extreme point of $\Psi\mathbb{B}_1^s$ is either a column of $\Psi$ or a column of $-\Psi$, and we get the result.

$2 \Rightarrow 1$ : If for all $v \in \mathrm{Ext}(\mathrm{im}(\Psi) \cap \mathbb{B}_p^r)$, $v$ or $-v$ is a column of $\Psi$, then $\mathrm{Ext}(\mathrm{im}(\Psi) \cap \mathbb{B}_p^r) \subseteq \Psi\mathbb{B}_1^s$. Taking convex hull, we obtain $\mathrm{im}(\Psi) \cap \mathbb{B}_p^r = \mathrm{conv}(\mathrm{Ext}(\mathrm{im}(\Psi) \cap \mathbb{B}_p^r)) \subseteq \Psi\mathbb{B}_1^s$. Since the converse inclusion always holds by Lemma 5.2, we obtain the equality.

$3 \iff 1$ : Follows from Lemma 5.2. □

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

**Lemma 5.4.** *Let $\Psi \in \mathbb{R}^{r\times s}$ 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,*

$$\frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)} = \min\left\{\beta \in \mathbb{R}_{\geq 1} : \mathrm{im}(\Psi^\top) \cap \mathbb{B}_\infty^s \subseteq \beta\Psi^\top\mathbb{B}_q^r\right\}. \tag{5.11}$$

*Proof.* We set $\beta^* := \min\left\{\beta \in \mathbb{R}_{\geq 1} : \mathrm{im}(\Psi^\top) \cap \mathbb{B}_\infty^s \subseteq \beta\Psi^\top\mathbb{B}_q^r\right\}$. Then

$$\begin{aligned}
\mathrm{im}(\Psi^\top) \cap \mathbb{B}_\infty^s \subseteq \beta^*\Psi^\top\mathbb{B}_q^r &\iff \left(\mathrm{im}(\Psi^\top) \cap \mathbb{B}_\infty^s\right)^\circ \supseteq \left(\beta^*\Psi^\top\mathbb{B}_q^r\right)^\circ \\
&\iff \mathrm{Null}(\Psi) + \mathbb{B}_1^s \supseteq \frac{1}{\beta^*}\Psi^{-1}\mathbb{B}_p^r \\
&\iff \mathrm{Null}(\Psi) + \beta^*\mathbb{B}_1^s \supseteq \Psi^{-1}\mathbb{B}_p^r \\
&\implies \beta^*\Psi\mathbb{B}_1^s \supseteq \Psi\Psi^{-1}\mathbb{B}_p^r = \mathrm{im}(\Psi) \cap \mathbb{B}_p^r.
\end{aligned}$$

Hence, we must have $\beta^* \geq \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)}$. Conversely, we have

$$\begin{aligned}
\mathrm{im}(\Psi) \cap \mathbb{B}_p^r \subseteq \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)}\Psi\mathbb{B}_1^s &\iff \left(\mathrm{im}(\Psi) \cap \mathbb{B}_p^r\right)^\circ \supseteq \left(\frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)}\Psi\mathbb{B}_1^s\right)^\circ \\
&\iff \mathrm{Null}(\Psi^\top) + \mathbb{B}_q^r \supseteq \mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)(\Psi^\top)^{-1}\mathbb{B}_\infty^s \\
&\iff \mathrm{Null}(\Psi^\top) + \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)}\mathbb{B}_q^r \supseteq (\Psi^\top)^{-1}\mathbb{B}_\infty^s \\
&\implies \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)}\Psi^\top\mathbb{B}_q^r \supseteq \Psi^\top(\Psi^\top)^{-1}\mathbb{B}_\infty^s = \mathrm{im}(\Psi^\top) \cap \mathbb{B}_\infty^s.
\end{aligned}$$

So, we also have the reverse inequality, $\beta^* \leq \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)}$. Hence, $\beta^* = \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}_1^s\right)}$. □

**Proof of Prop. 1.**

*Proof.* By Lemma 5.4, we have $\operatorname{im}(\Psi^\top) \cap \mathbb{B}^s_\infty \subseteq \frac{1}{\mathfrak{r}_p(\Psi\mathbb{B}^s_1)}\Psi^\top\mathbb{B}^r_q$. Hence,

$$\begin{aligned} \max_{w\in\operatorname{im}(\Psi^\top)\cap\mathbb{B}^s_\infty} \bar{y}^\top w &\leq \max_{w\in\frac{1}{\mathfrak{r}_p(\Psi\mathbb{B}^s_1)}\Psi^\top\mathbb{B}^r_q} \bar{y}^\top w \qquad (5.12)\\ &= \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}^s_1\right)} \max_{w\in\Psi^\top\mathbb{B}^r_q} \bar{y}^\top w \\ &= \frac{1}{\mathfrak{r}_p\left(\Psi\mathbb{B}^s_1\right)} \max_{e\in\mathbb{B}^r_q} \bar{y}^\top\Psi^\top e \\ &= \frac{\|\Psi\bar{y}\|_p}{\mathfrak{r}_p\left(\Psi\mathbb{B}^s_1\right)}. \end{aligned}$$

If $p \in \{1, \infty\}$ and $\mathfrak{r}_p\left(\Psi\mathbb{B}^s_1\right) = 1$ (i.e. $\Psi$ is perfectly expressive), then, by Lemma 5.3, equality holds in (5.12) with $\mathfrak{r}_p\left(\Psi\mathbb{B}^s_1\right) = 1$, and the result follows. □

# References

[1] Ralph Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, 1970.