A Bipartite Maximization Guarantees

Lemma 3.2. For any real \( m \times n \), rank-\( r \) matrix \( \tilde{A} \) and arbitrary norm-bounded sets \( \mathcal{X} \subset \mathbb{R}^{m \times k} \) and \( \mathcal{Y} \subset \mathbb{R}^{n \times k} \), let

\[
(\tilde{X}_*, \tilde{Y}_*) \triangleq \arg \max_{X \in \mathcal{X}, Y \in \mathcal{Y}} \text{Tr}(X^\top \tilde{A} Y).
\]

If there exist operators \( P_X : \mathbb{R}^{m \times k} \to \mathcal{X} \) such that \( P_X(L) = \arg \max_{X \in \mathcal{X}} \text{Tr}(X^\top L) \)

and similarly, \( P_Y : \mathbb{R}^{n \times k} \to \mathcal{Y} \) such that

\[
P_Y(R) = \arg \max_{Y \in \mathcal{Y}} \text{Tr}(R^\top Y)
\]

with running times \( T_X \) and \( T_Y \), respectively, then Algorithm 1 outputs \( \tilde{X} \in \mathcal{X} \) and \( \tilde{Y} \in \mathcal{Y} \) such that

\[
\text{Tr}(\tilde{X}^\top \tilde{A} \tilde{Y}) \geq \text{Tr}(\tilde{X}_*^\top \tilde{A} \tilde{Y}_*) - 2\epsilon \sqrt{k} \cdot \|\tilde{A}\|_2 \cdot \mu_X \cdot \mu_Y,
\]

where \( \mu_X \triangleq \max_{X \in \mathcal{X}} \|X\|_F \) and \( \mu_Y \triangleq \max_{Y \in \mathcal{Y}} \|Y\|_F \).

Proof. In the sequel, \( \tilde{U} \), \( \tilde{\Sigma} \) and \( \tilde{V} \) are used to denote the \( \epsilon \)-truncated singular value decomposition of \( \tilde{A} \).

Without loss of generality, we assume that \( \mu_X = \mu_Y = 1 \) since the variables in \( \mathcal{X} \) and \( \mathcal{Y} \) can be normalized by \( \mu_X \) and \( \mu_Y \), respectively, while simultaneously scaling the singular values of \( \tilde{A} \) by a factor of \( \mu_X \cdot \mu_Y \).

Then \( \|Y\|_{\infty,2} \leq 1, \forall Y \in \mathcal{Y} \), where \( \|Y\|_{\infty,2} \) denotes the maximum of the \( \ell_2 \)-norm of the columns of \( Y \).

Let \( \tilde{X}_*, \tilde{Y}_* \) be the optimal pair on \( \tilde{A} \), i.e.,

\[
(\tilde{X}_*, \tilde{Y}_*) \triangleq \arg \max_{X \in \mathcal{X}, Y \in \mathcal{Y}} \text{Tr}(X^\top \tilde{A} Y)
\]

and define the \( r \times k \) matrix \( \tilde{C}_* \triangleq \tilde{V}^\top \tilde{Y}_* \). Note that

\[
\|\tilde{C}_*\|_{\infty,2} = \|\tilde{V}^\top \tilde{Y}_*\|_{\infty,2} = \max_{1 \leq i \leq k} \|\tilde{V}^\top [\tilde{Y}_*]_i\|_2 \leq 1,
\]

with the last inequality following from the facts that \( \|Y\|_{\infty,2} \leq 1, \forall Y \in \mathcal{Y} \) and the columns of \( \tilde{V} \) are orthonormal. Alg. 1 iterates over the points in \( (\mathbb{B}_2^{-1})^\otimes k \).

The latter is used to describe the set of \( r \times k \) matrices whose columns have \( \ell_2 \) norm at most equal to 1. At each point, the algorithm computes a candidate solution.

By (14), the \( \epsilon \)-net contains an \( r \times k \) matrix \( C \) such that

\[
\|C - \tilde{C}_*\|_{\infty,2} \leq \epsilon.
\]

Let \( X_*, Y_* \) be the candidate pair computed at \( C \) by the two step maximization, i.e.,

\[
X_* \triangleq \arg \max_{X \in \mathcal{X}} \text{Tr}(X^\top \tilde{U} \tilde{\Sigma} C)
\]

and

\[
Y_* \triangleq \arg \max_{Y \in \mathcal{Y}} \text{Tr}(X_*^\top \tilde{A} Y).
\]

We show that the objective values achieved by the candidate pair \( X_*, Y_* \) satisfies the inequality of the lemma implying the desired result.

By the definition of \( \tilde{C}_* \) and the linearity of the trace,

\[
\text{Tr}(\tilde{X}_*^\top \tilde{A} \tilde{Y}_*)
\]

\[
= \text{Tr}(\tilde{X}_*^\top \tilde{U} \tilde{\Sigma} \tilde{C}_*)
\]

\[
= \text{Tr}(\tilde{X}_*^\top \tilde{U} \tilde{\Sigma} \tilde{C}_*) + \text{Tr}(\tilde{X}_*^\top \tilde{U} \tilde{\Sigma} (\tilde{C}_* - \tilde{C}))
\]

\[
\leq \text{Tr}(\tilde{X}_*^\top \tilde{U} \tilde{\Sigma} \tilde{C}_*) + \text{Tr}(\tilde{X}_*^\top \tilde{U} \tilde{\Sigma} (\tilde{C}_* - \tilde{C})).
\]

The inequality follows from the fact that (by definition (15)) \( X_* \) maximizes the first term over all \( X \in \mathcal{X} \).

We compute an upper bound on the right hand side of (16). Define

\[
\bar{Y} \triangleq \arg \min_{Y \in \mathcal{Y}} \|\bar{V}^\top Y - C - \tilde{C}\|_{\infty,2}.
\]

(We note that \( \bar{Y} \) is used for the analysis and is never explicitly calculated.) Further, define the \( r \times k \) matrix \( \tilde{C} \triangleq \tilde{V}^\top \bar{Y} \). By the linearity of the trace operator

\[
\text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} C_\dagger)
\]

\[
= \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} \tilde{C}_\dagger) + \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} (C_\dagger - \tilde{C}))
\]

\[
= \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} \tilde{Y}_\dagger) + \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} (C_\dagger - \tilde{C}))
\]

\[
\leq \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} \bar{Y} + \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} (C_\dagger - \tilde{C}))
\]

\[
= \text{Tr}(X_\dagger^\top \tilde{A} \bar{Y} + \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} (C_\dagger - \tilde{C})).
\]

(17)

The inequality follows from the fact that (by definition (15)) \( Y_\dagger \) maximizes the first term over all \( Y \in \mathcal{Y} \).

Combining (17) and (16), and rearranging the terms, we obtain

\[
\text{Tr}(X_\dagger^\top \tilde{A} \tilde{Y}_\dagger) - \text{Tr}(X_\dagger^\top \tilde{A} \bar{Y}_\dagger)
\]

\[
\leq \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} (\tilde{C}_\dagger - \bar{C}_\dagger)) + \text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} (C_\dagger - \tilde{C})).
\]

(18)

By Lemma C.10,

\[
\text{Tr}(X_\dagger^\top \tilde{U} \tilde{\Sigma} (C_\dagger - \tilde{C}))
\]

\[
\leq \|X_\dagger^\top \tilde{U}\|_F \cdot \|\tilde{\Sigma}\|_2 \cdot \|\tilde{C}_\dagger - \tilde{C}\|_F
\]

\[
\leq \|X_\dagger\|_F \cdot \sigma_1(\tilde{A}) \cdot \sqrt{k} \cdot \epsilon
\]

\[
\leq \max_{X \in \mathcal{X}} \|X\|_F \cdot \sigma_1(\tilde{A}) \cdot \sqrt{k} \cdot \epsilon
\]

\[
\leq \sigma_1(\tilde{A}) \cdot \sqrt{k} \cdot \epsilon.
\]

(19)
Similarly,
\[
\| \text{Tr}(X^T_{\sigma} U \Sigma (C_{\sigma} - \hat{C})) \| \leq \| X_{\sigma} U \|_F \cdot \| \Sigma \|_2 \cdot \| C_{\sigma} - \hat{C} \|_F
\]
\[
\leq \max_{X \in \mathcal{X}} \| X \|_F \cdot \sigma_1(A) \cdot \sqrt{k} \cdot \epsilon
\]
\[
\leq \sigma_1(A) \cdot \sqrt{k} \cdot \epsilon.
\]  
(20)

The second inequality follows from the fact that by the definition of \( C \),
\[
\| \hat{C} - C_{\sigma} \|_{\infty, 2} = \| \tilde{V}^T \tilde{Y} - C_{\sigma} \|_{\infty, 2} \leq \| \tilde{V}^T \tilde{Y} - C_{\sigma} \|_{\infty, 2} = \| \hat{C} - C_{\sigma} \|_{\infty, 2} \leq \epsilon,
\]
which implies that
\[
\| \hat{C} - C_{\sigma} \|_F \leq \sqrt{k} \cdot \epsilon.
\]
Continuing from (18) under (19) and (20),
\[
\text{Tr}(X^T_{\sigma} A Y_{\sigma}) \geq \text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}_*) - 2 \cdot \epsilon \cdot \sqrt{k} \cdot \sigma_1(A).
\]
Recalling that the singular values of \( \hat{A} \) have been scaled by a factor of \( \mu_x \cdot \mu_y \) yields the desired result.

The runtime of Alg. 1 follows from the cost per iteration and the cardinality of the \( \epsilon \)-net. Matrix multiplications can exploit the truncated singular value decomposition of \( \hat{A} \) which is performed only once.

**Lemma A.6.** For any \( A, \tilde{A} \in \mathbb{R}^{m \times n} \), and norm-bounded sets \( \mathcal{X} \subseteq \mathbb{R}^{m \times k} \) and \( \mathcal{Y} \subseteq \mathbb{R}^{n \times k} \), let
\[
(X_*, Y_*) \triangleq \arg \max_{X \in \mathcal{X}, Y \in \mathcal{Y}} \text{Tr}(X^T A Y),
\]
and
\[
(\tilde{X}_*, \tilde{Y}_*) \triangleq \arg \max_{X \in \mathcal{X}, Y \in \mathcal{Y}} \text{Tr}(X^T \tilde{A} Y).
\]
For any \((\tilde{X}, \tilde{Y}) \in \mathcal{X} \times \mathcal{Y}\) such that
\[
\text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) \geq \gamma \cdot \text{Tr}(\tilde{X}_*^T \tilde{A} \tilde{Y}_*) - C
\]
for some \(0 < \gamma \leq 1\), we have
\[
\text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) \geq \gamma \cdot \text{Tr}(X_*^T A Y_*) - C
\]
\[
- 2 \cdot \| A - \tilde{A} \|_2 \cdot \mu_x \cdot \mu_y,
\]
where \( \mu_x \triangleq \max_{X \in \mathcal{X}} \| X \|_F \) and \( \mu_y \triangleq \max_{Y \in \mathcal{Y}} \| Y \|_F \).

**Proof.** By the optimality of \( \tilde{X}_*, \tilde{Y}_* \) for \( \tilde{A} \), we have
\[
\text{Tr}(\tilde{X}_*^T \tilde{A} \tilde{Y}_*) \geq \text{Tr}(X_*^T A Y_*)
\]
In turn, for any \((\tilde{X}, \tilde{Y}) \in \mathcal{X} \times \mathcal{Y}\) such that
\[
\text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) \geq \gamma \cdot \text{Tr}(\tilde{X}_*^T \tilde{A} \tilde{Y}_*) - C
\]
for some \(0 < \gamma < 1\) (if such pairs exist), we have
\[
\text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) \geq \gamma \cdot \text{Tr}(X_*^T A Y_*) - C.
\]

By the linearity of the trace operator,
\[
\text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y})
\]
\[
= \text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) - \text{Tr}(\tilde{X}^T (A - \tilde{A}) \tilde{Y})
\]
\[
\leq \text{Tr}(X_*^T A Y_*) + |\text{Tr}(\tilde{X}^T (A - \tilde{A}) \tilde{Y})|.
\]
(22)

By Lemma C.10,
\[
|\text{Tr}(\tilde{X}^T (A - \tilde{A}) \tilde{Y})|
\]
\[
\leq \| \tilde{X} \|_F \cdot \| \tilde{Y} \|_F \cdot \| A - \tilde{A} \|_2
\]
\[
\leq \| A - \tilde{A} \|_2 \cdot \max_{X \in \mathcal{X}} \| X \|_F \cdot \max_{Y \in \mathcal{Y}} \| Y \|_F \triangleq R.
\]
(23)

Continuing from (22),
\[
\text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) \leq \text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) + R.
\]
(24)

Similarly,
\[
\text{Tr}(X_*^T A Y_*)
\]
\[
= \text{Tr}(X_*^T A Y_*) - \text{Tr}(X_*^T (A - \tilde{A}) \tilde{Y}_*)
\]
\[
\geq \text{Tr}(X_*^T A Y_*) - |\text{Tr}(X_*^T (A - \tilde{A}) \tilde{Y}_*)|
\]
\[
\geq \text{Tr}(X_*^T A Y_*) - R.
\]
(25)

Combining the above, we have
\[
\text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) \geq \text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y}) - R
\]
\[
\geq \gamma \cdot \text{Tr}(X_*^T A Y_*) - R - C
\]
\[
\geq \gamma \cdot (\text{Tr}(X_*^T A Y_*) - R) - R - C
\]
\[
= \gamma \cdot \text{Tr}(X_*^T A Y_*) - (1 + \gamma) \cdot R - C
\]
\[
\geq \gamma \cdot \text{Tr}(X_*^T A Y_*) - 2 \cdot R - C,
\]
where the first inequality follows from (24) the second from (21), the third from (25), and the last from the fact that \( R \geq 0 \). This concludes the proof.

**Lemma 3.3.** For any \( A \in \mathbb{R}^{m \times n} \), let
\[
(X_*, Y_*) \triangleq \arg \max_{X \in \mathcal{X}, Y \in \mathcal{Y}} \text{Tr}(X^T A Y),
\]
where \( \mathcal{X} \subseteq \mathbb{R}^{m \times k} \) and \( \mathcal{Y} \subseteq \mathbb{R}^{n \times k} \) are sets satisfying the conditions of Lemma 3.2. Let \( \tilde{A} \) be a rank-\( r \) approximation of \( A \), and \( \tilde{X} \in \mathcal{X}, \tilde{Y} \in \mathcal{Y} \) be the output of Alg. 1 with input \( A \) and accuracy \( \epsilon \). Then,
\[
\text{Tr}(X_*^T A Y_*) - \text{Tr}(\tilde{X}^T \tilde{A} \tilde{Y})
\]
\[
\leq 2 \cdot (\epsilon \sqrt{k} \cdot \| \tilde{A} \|_2 + \| A - \tilde{A} \|_2) \cdot \mu_x \cdot \mu_y,
\]
where \( \mu_X \triangleq \max_{X \in \mathcal{X}} \| X \|_F \) and \( \mu_Y \triangleq \max_{Y \in \mathcal{Y}} \| Y \|_F \).

**Proof.** The proof follows the approximation guarantees of Alg. 1 in Lemma 3.2 and Lemma A.6.
B Correctness of Algorithm 2

In the sequel, we use \( \|X\|_{\infty,1} \) to denote the maximum of the \( \ell_1 \) norm of the rows of \( X \). When \( X \in \{0,1\}^{d \times k} \), the constraint \( \|X\|_{\infty,1} = 1 \) effectively implies that each row of \( X \) has exactly one nonzero entry.

**Lemma C.7.** Let \( \mathcal{X} \triangleq \{X \in \{0,1\}^{d \times k} : \|X\|_{\infty,1} = 1\} \).

For any \( d \times k \) real matrix \( X \), Algorithm 2 outputs

\[
\hat{X} = \arg \max_{X \in \mathcal{X}} \text{Tr}(X^T L),
\]

in time \( O(k \cdot d) \)

**Proof.** By construction, each row of \( X \) has exactly one nonzero entry. Let \( j_i \in [k] \) denote the index of the nonzero entry in the \( i \)-th row of \( X \). For any \( X \in \mathcal{X} \),

\[
\text{Tr}(X^T L) = \sum_{j=1}^{k} X_{ji}^T 1_j = \sum_{j=1}^{k} \sum_{i \in \supp(x_j)} 1 \cdot L_{ij} \leq \sum_{i=1}^{d} L_{ij} \leq \sum_{i=1}^{d} \max_{j \in [k]} L_{ij}. \tag{26}
\]

Algorithm 2 achieves equality in (26) due to the choice of \( j_i \) in line 3. Finally, the running time follows immediately from the \( O(k) \) time required to determine the maximum entry of each of the \( d \) rows of \( L \).

C Auxiliary Lemmas

**Lemma C.8.** For any \( A, B \in \mathbb{R}^{n \times k} \),

\[
|\langle A, B \rangle| \triangleq |\text{Tr}(A^T B)| \leq \|A\|_F \cdot \|B\|_F.
\]

**Proof.** Treating \( A \) and \( B \) as vectors, the lemma follows immediately from Lemma C.7 for \( p = q = 2 \).

**Lemma C.9.** For any two real matrices \( A \) and \( B \) of appropriate dimensions,

\[
\|AB\|_F \leq \min\{\|A\|_2 \cdot \|B\|_F, \|A\|_F \cdot \|B\|_2\}.
\]

**Proof.** Let \( b_i \) denote the \( i \)-th column of \( B \). Then,

\[
\|AB\|_F^2 = \sum_i \|Ab_i\|_2^2 \leq \sum_i \|A\|_2^2 \cdot \|b_i\|_2^2 = \|A\|_2^2 \cdot \|B\|_2^2.
\]

Similarly, using the previous inequality,

\[
\|AB\|_2^2 = \|B^T A^T\|_F^2 \leq \|B\|_2^2 \cdot \|A^T\|_2^2 = \|B\|_2^2 \cdot \|A\|_2^2.
\]

The desired result follows combining the two upper bounds.

**Lemma C.10.** For any real \( m \times k \) matrix \( X \), \( mn \) matrix \( A \), and \( n \times k \) matrix \( Y \),

\[
|\text{Tr}(X^T AY)| \leq \|X\|_F \cdot \|A\|_2 \cdot \|Y\|_F.
\]

**Proof.** We have

\[
|\text{Tr}(X^T AY)| \leq \|X\|_F \cdot \|AY\|_F \leq \|X\|_F \cdot \|A\|_2 \cdot \|Y\|_F,
\]

with the first inequality following from Lemma C.8 on \( |\langle X, AY \rangle| \) and the second from Lemma C.9.

**Lemma C.11.** For any real \( m \times n \) matrix \( A \), and pair of \( m \times k \) matrix \( X \) and \( n \times k \) matrix \( Y \) such that \( X^T X = I_k \) and \( Y^T Y = I_k \) with \( k \leq \min\{m, n\} \), the following holds:

\[
|\text{Tr}(X^T AY)| \leq \sqrt{k} \cdot (\sum_{i=1}^{k} \sigma_i^2(A))^{1/2}.
\]

**Proof.** By Lemma C.8,

\[
|\langle X, AY \rangle| = |\text{Tr}(X^T AY)| \leq \|X\|_F \cdot \|AY\|_F = \sqrt{k} \cdot \|AY\|_F.
\]

where the last inequality follows from the fact that \( \|X\|_F^2 = \text{Tr}(X^T X) = \text{Tr}(I_k) = k \). Further, for any \( Y \) such that \( Y^T Y = I_k \),

\[
\|AY\|_F^2 \leq \max_{Y \in \mathbb{R}^{n \times k}} \|AY\|_F^2 = \sum_{i=1}^{k} \sigma_i^2(A). \tag{27}
\]

Combining the two inequalities, the result follows.

**Lemma C.12.** For any real \( m \times n \) matrix \( A \), and any \( k \leq \min\{m, n\} \),

\[
\max_{Y \in \mathbb{R}^{n \times k}} \|AY\|_F = \left(\sum_{i=1}^{k} \sigma_i^2(A)\right)^{1/2}.
\]

The above equality is realized when the \( k \) columns of \( Y \) coincide with the \( k \) leading right singular vectors of \( A \).
Lemma C.13. Finally, it is straightforward to verify that if

\[ \|AY\|_F^2 = \|U\Sigma V^T Y\|_F^2 = \|\Sigma V^T Y\|_F^2. \]  

(28)

Continuing from (28),

\[
\|\Sigma V^T Y\|_F^2 = \text{Tr}(Y^T V \Sigma^2 V^T Y) \\
= \sum_{i=1}^{k} y_i^T \left( \sum_{j=1}^{d} \sigma_j^2 \cdot v_j v_j^T \right) y_i \\
= \sum_{j=1}^{d} \sigma_j^2 \cdot \sum_{i=1}^{k} (v_j^T y_i)^2.
\]

Let \( z_j \triangleq \sum_{i=1}^{k} (v_j^T y_i)^2, \ j = 1, \ldots, d. \) Note that each individual \( z_j \) satisfies

\[ 0 \leq z_j = \sum_{i=1}^{k} (v_j^T y_i)^2 \leq \|v_j\|^2 = 1, \]

where the last inequality follows from the fact that the columns of \( Y \) are orthonormal. Further,

\[
\sum_{j=1}^{d} z_j = \sum_{j=1}^{k} \sum_{i=1}^{d} (v_j^T y_i)^2 = \sum_{i=1}^{k} \sum_{j=1}^{d} (v_j^T y_i)^2 \\
= \sum_{i=1}^{k} \|y_i\|^2 = k.
\]

Combining the above, we conclude that

\[ \|AY\|_F^2 = \sum_{j=1}^{d} \sigma_j^2 \cdot z_j \leq \sigma_1^2 + \ldots + \sigma_k^2. \]  

(29)

Finally, it is straightforward to verify that if \( y_i = v_i, \ i = 1, \ldots, k, \) then (29) holds with equality. \( \square \)

Lemma C.13. For any real \( m \times n \) matrix \( A, \) let \( \sigma_i(A) \) be the \( i \)th largest singular value. For any \( r, k \leq \min\{m, n\}, \)

\[ \sum_{i=r+1}^{r+k} \sigma_i(A) \leq \frac{k}{\sqrt{r}+k} \|A\|_F. \]  

Proof. By the Cauchy-Schwartz inequality,

\[
\sum_{i=r+1}^{r+k} \sigma_i(A) = \sum_{i=r+1}^{r+k} |\sigma_i(A)| \leq \left( \sum_{i=r+1}^{r+k} \sigma_i^2(A) \right)^{1/2} \|1_k\|_2 \\
= \sqrt{k} \cdot \left( \sum_{i=r+1}^{r+k} \sigma_i^2(A) \right)^{1/2}.
\]

Note that \( \sigma_{r+1}(A), \ldots, \sigma_{r+k}(A) \) are the \( k \) smallest among the \( r+k \) largest singular values. Hence,

\[
\sum_{i=r+1}^{r+k} \sigma_i^2(A) \leq \frac{k}{r+k} \sum_{i=1}^{r+k} \sigma_i^2(A) \leq \frac{k}{r+k} \sum_{i=1}^{l} \sigma_i^2(A) \\
= \frac{k}{r+k} \|A\|_F^2.
\]

Combining the two inequalities, the desired result follows. \( \square \)

Corollary 1. For any real \( m \times n \) matrix \( A, \) the \( r \)th largest singular value \( \sigma_r(A) \) satisfies \( \sigma_r(A) \leq \|A\|_F/\sqrt{r}. \)

Proof. It follows immediately from Lemma C.13. \( \square \)

First, we define the \( \| \cdot \|_{\infty,2} \) norm of a matrix as the \( l_2 \) norm of the column with the maximum \( l_2 \) norm, i.e., for an \( r \times k \) matrix \( C \)

\[ \|C\|_{\infty,2} = \max_{1 \leq i \leq k} \|c_i\|_2. \]

Note that

\[
\|C\|_F^2 = \sum_{i=1}^{k} \|c_i\|_2^2 \leq k \cdot \max_{1 \leq i \leq k} \|c_i\|_2^2 \\
= k \cdot \left( \max_{1 \leq i \leq k} \|c_i\|_2 \right)^2 = k \cdot \|C\|_{\infty,2}. \]  

(30)