I. APPENDIX

A. Proof of Auxiliary Lemmas for Theorem 1

Our proof follows similar arguments as [1] with necessary modifications for beamspace and multi-snapshot scenario. For completeness, we provide all auxiliary lemmas used.

Preliminaries

Let $\mathbf{S}_1, \mathbf{S}_2$ be any orthonormal bases for $\mathcal{R}(\mathbf{U}_y)$ and $\mathcal{R}(\widehat{\mathbf{U}}_y)$, respectively. The principal (or canonical) angles between the subspaces $\mathcal{R}(\mathbf{U}_y)$ and $\mathcal{R}(\widehat{\mathbf{U}}_y)$ are defined as the $\mathbf{\Theta}(\mathbf{S}_1, \mathbf{S}_2) := [\omega_1, \omega_2, \cdots, \omega_S]^T$ where $\omega_k \in [0, \pi/2]$ satisfies:

$$\cos(\omega_i) = \sigma_i(\mathbf{S}_1^H \mathbf{S}_2) \tag{48}$$

We consider the SVD of $\mathbf{S}_1^H \mathbf{S}_2 = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}}\widetilde{\mathbf{V}}^H$. Since ESPRIT is invariant to the exact choice of the basis, for our analysis we will consider the orthonormal bases for $\mathcal{R}(\mathbf{U}_y)$ and $\mathcal{R}(\widehat{\mathbf{U}}_y)$ as $\mathbf{U}_y = \mathbf{S}_1\widetilde{\mathbf{U}}$, and $\widehat{\mathbf{U}}_y = \mathbf{S}_2\widetilde{\mathbf{V}}$. In this case, it can be verified that the principal angles defined in (48) can be written as:

$$\cos(\omega_i) = |\mathbf{u}_i^H \widehat{\mathbf{u}}_i|$$

Here we assumed that the singular vectors are ordered such that $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_S$. We also denote

$$\sin(\mathbf{\Theta}(\mathbf{U}_y, \widehat{\mathbf{U}}_y)) := [\sin(\omega_1), \sin(\omega_2), \cdots, \sin(\omega_S)]^T$$

The augmented noise matrix is given by:

$$\mathbf{N}_s := \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{bmatrix}$$

where $\mathbf{N}_1, \mathbf{N}_2 \in \mathbb{C}^{M-1\times T}$ represent matrices formed by selecting the first $M-1$ rows and last $M-1$ rows of $\mathbf{N}$, respectively. Let $\widetilde{\mathbf{N}} = \mathbf{W}^H \mathbf{N}_s$, we have the following bound:

$$\begin{aligned} \|\widetilde{\mathbf{N}}\|_2^2 &\leq \|\mathbf{W}\|_2^2(\|\mathbf{N}_1\|_2^2 + \|\mathbf{N}_2\|_2^2) \\ &\leq 2\|\mathbf{W}\|_2^2\|\mathbf{N}\|_2^2 \end{aligned} \tag{49}$$

where the first inequality follows from the fact that $|\mathbf{N}_s\|_2^2 \leq \|\mathbf{N}_1\|_2^2 + \|\mathbf{N}_2\|_2^2$, and the second inequality holds since both $\mathbf{N}_1, \mathbf{N}_2$ are submatrices of $\mathbf{N}$.

For any matrix $\mathbf{F}$, we adopt the notation $\sigma_{\max}(\mathbf{F}) := \|\mathbf{F}\|_2$, and $\sigma_{\min}(\mathbf{F}) := 1/\|\mathbf{F}^\dagger\|_2$. We first use Wedin's theorem [2] to bound $\|\mathbf{U}_y - \widehat{\mathbf{U}}_y\|_2$.

Lemma 1 (Wedin's Theorem [2])**.** *Consider matrices* $\mathbf{A}, \mathbf{B}, \mathbf{N} \in \mathbb{C}^{M\times N}$ *such that*

$$\mathbf{B} = \mathbf{A} + \mathbf{N}$$

Consider the Singular Value Decompositions of $\mathbf{A}$ *and* $\mathbf{B}$*:*

$$\mathbf{A} = [\mathbf{U}_1\ \mathbf{U}_0] \begin{bmatrix} \mathbf{\Sigma_1} & \\ & \mathbf{\Sigma_0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_0 \end{bmatrix}^H$$

$$\mathbf{B} = [\widetilde{\mathbf{U}}_1\ \widetilde{\mathbf{U}}_0] \begin{bmatrix} \widetilde{\mathbf{\Sigma}}_1 & \\ & \widetilde{\mathbf{\Sigma}}_0 \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{V}}_1 \\ \widetilde{\mathbf{V}}_0 \end{bmatrix}^H$$

where $\mathbf{U}_1 \in \mathbb{C}^{M\times L}$, $\widetilde{\mathbf{U}}_1 \in \mathbb{C}^{M\times L}$ *consist of the* L *principal singular vectors of* $\mathbf{A}$ *and* $\mathbf{B}$*, respectively. Define* $\mathbf{A}_1 := \mathbf{U}_1\mathbf{\Sigma_1}\mathbf{V}_1^H$, $\mathbf{A}_0 := \mathbf{U}_0\mathbf{\Sigma_0}\mathbf{V}_0^H$, $\mathbf{B}_1 := \widetilde{\mathbf{U}}_1\widetilde{\mathbf{\Sigma}}_1\widetilde{\mathbf{V}}_1^H$, $\mathbf{B}_0 := \widetilde{\mathbf{U}}_0\widetilde{\mathbf{\Sigma}}_0\widetilde{\mathbf{V}}_0^H$*.If* $\sigma_{max}(\mathbf{A}_0) \leq \alpha$ *and* $\sigma_{min}(\mathbf{B}_1) \geq \alpha + \delta$ *for some* $\alpha \geq 0$ *and* $\delta > 0$*, the following holds*

$$\|\sin\mathbf{\Theta}(\mathcal{R}(\mathbf{A}_1), \mathcal{R}(\mathbf{B}_1))\|_\infty \leq \frac{\max\{\|\mathbf{N}\mathbf{V}_1\|_2, \|\mathbf{N}^H\mathbf{U}_1\|_2\}}{\delta}$$

Lemma 2. *Consider the matrices* $\mathbf{A}, \mathbf{B}_1, \mathbf{U}_1, \mathbf{V}_1$ *defined in Lemma 1. If rank*$(\mathbf{A}) = L$*, and* $\|\mathbf{N}\|_2 \leq \sigma_L(\mathbf{A})/2$*, the following holds*

$$\|\sin\mathbf{\Theta}(\mathcal{R}(\mathbf{A}), \mathcal{R}(\mathbf{B}_1))\|_\infty \leq \frac{2\max\{\|\mathbf{N}\mathbf{V}_1\|_2, \|\mathbf{N}^H\mathbf{U}_1\|_2\}}{\sigma_L(\mathbf{A})}$$

Proof. Note that since rank$(\mathbf{A}) = L$, we have $\mathbf{A}_0 = \mathbf{0}$, and $\sigma_{\min}(\mathbf{A}) = \sigma_L(\mathbf{A})$. Using Weyl's theorem [3] for matrix perturbation, we can write

$$\sigma_{\min}(\mathbf{B}_1) \geq \sigma_{\min}(\mathbf{A}) - \|\mathbf{N}\|_2 \geq \frac{\sigma_{\mathrm{L}}(\mathbf{A})}{2}$$

where the last inequality follows from the assumption $\|\mathbf{N}\|_2 \leq \sigma_L(\mathbf{A})/2$. The conditions of Lemma 1 are satisfied with $\alpha = 0$ and $\delta = \sigma_L(\mathbf{A})$ completing the proof of Lemma 2. □

We will also be using the following standard result from [4, Pg. 36].

Lemma 3. *For any matrices* $\mathbf{A} \in \mathbb{C}^{M\times K}$*, and* $\mathbf{B} \in \mathbb{C}^{K\times T}$*, $(M > K)$ where rank*$(\mathbf{A}) = K$*, we have*

$$\sigma_K(\mathbf{AB}) \geq \sigma_K(\mathbf{A})\sigma_K(\mathbf{B})$$

Lemma 4. *Let* $\widehat{\mathbf{Y}} = \mathbf{BX} + \widetilde{\mathbf{N}}$*, where Rank* $(\mathbf{BX}) = S$*. Consider the Singular Value Decompositions:* $\mathbf{BX} = \mathbf{U}_y\mathbf{\Sigma}_y\mathbf{V}_y^H$*,* $\widehat{\mathbf{Y}} = [\widehat{\mathbf{U}}_y\ \ \widehat{\mathbf{U}}_n]\widehat{\mathbf{\Sigma}}_y[\widehat{\mathbf{V}}_y^H\ \ \widehat{\mathbf{V}}_n^H]^H$*, where* $\mathbf{U}_y, \widehat{\mathbf{U}}_y \in \mathbb{C}^{2N\times S}$ *consists of the* S *principle singular vectors. Assuming that the noise is bounded as* $\|\widetilde{\mathbf{N}}\|_2 \leq \sigma_S(\mathbf{B})\sigma_S(\mathbf{X})/2$*, the following holds*

$$\|\mathbf{U}_y - \widehat{\mathbf{U}}_y\|_2 \leq \frac{2\sqrt{2S}\|\widetilde{\mathbf{N}}\|_2}{\sigma_S(\mathbf{B})\sigma_S(\mathbf{X})} \tag{50}$$

Proof. When the noise $\widetilde{\mathbf{N}}$ is bounded by $\|\widetilde{\mathbf{N}}\|_2 \leq \sigma_S(\mathbf{B})\sigma_S(\mathbf{X})/2 \leq \sigma_S(\mathbf{BX})/2$, the assumptions of Lemma 2 are satisfied for $L = S$, which implies

$$\begin{aligned} \|\sin\mathbf{\Theta}(\mathbf{U}_y, \widehat{\mathbf{U}}_y)\|_\infty &\leq \frac{2\max\{\|\widetilde{\mathbf{N}}\mathbf{V}_y\|_2, \|\widetilde{\mathbf{N}}^H\mathbf{U}_y\|_2\}}{\sigma_S(\mathbf{BX})} \\ &\leq \frac{2\max\{\|\widetilde{\mathbf{N}}\mathbf{V}_y\|_2, \|\widetilde{\mathbf{N}}^H\mathbf{U}_y\|_2\}}{\sigma_S(\mathbf{B})\sigma_S(\mathbf{X})} \end{aligned}$$

Using the fact $\|\mathbf{V}_y\|_2 = 1$, $\|\mathbf{U}_y\|_2 = 1$, we have

$$\|\sin\mathbf{\Theta}(\mathbf{U}_y, \widehat{\mathbf{U}}_y)\|_\infty \leq \frac{2\|\widetilde{\mathbf{N}}\|_2}{\sigma_S(\mathbf{B})\sigma_S(\mathbf{X})} \tag{51}$$

Now, under the canonical basis assumption, we have $\|\sin\mathbf{\Theta}(\mathbf{U}_y, \widehat{\mathbf{U}}_y)\|_\infty = \sin(\omega_1)$ and for $i = 1, 2, \cdots S$

$$\|\widehat{\mathbf{u}}_i - \mathbf{u}_i\|_2^2 = 2(1 - \cos\omega_k) \leq 2(1 - \cos^2\omega_k) \leq 2\sin^2\omega_k$$

Therefore,

$$\begin{aligned} \|\mathbf{U}_y - \widehat{\mathbf{U}}_y\|_2 \leq \|\mathbf{U}_y - \widehat{\mathbf{U}}_y\|_F = \left(\sum_{i=1}^{S} \|\widehat{\mathbf{u}}_i - \mathbf{u}_i\|_2^2\right)^{1/2} \\ \leq (2S\sin^2\omega_1)^{1/2} = \sqrt{2S}\sin\omega_1 \end{aligned} \tag{52}$$

The proof is completed by combining (52) and (51). $\square$

Lemma 5. *Consider the measurement model in (14). If rank*$(\mathbf{BX}) = S$*, and* $\|\mathbf{U}_y - \widehat{\mathbf{U}}_y\|_2 \le \sigma_S(\mathbf{U}_1)/2$*, then*

$$\|\mathbf{\Psi} - \widehat{\mathbf{\Psi}}\|_2 \le \frac{7\|\mathbf{U}_y - \widehat{\mathbf{U}}_y\|_2}{\sigma_S(\mathbf{U}_1)^2} \tag{53}$$

Proof. Notice that

$$\begin{aligned}
\|\mathbf{\Psi} - \widehat{\mathbf{\Psi}}\|_2 &= \|(\widehat{\mathbf{U}}_1^\dagger - \mathbf{U}_1^\dagger)\widehat{\mathbf{U}}_2 + \mathbf{U}_1^\dagger(\widehat{\mathbf{U}}_2 - \mathbf{U}_2)\|_2 \\
&\le \|(\widehat{\mathbf{U}}_1^\dagger - \mathbf{U}_1^\dagger)\|_2\|\widehat{\mathbf{U}}_2\|_2 + \|\mathbf{U}_1^\dagger\|_2\|(\widehat{\mathbf{U}}_2 - \mathbf{U}_2)\|_2 \\
&\le \|(\widehat{\mathbf{U}}_1^\dagger - \mathbf{U}_1^\dagger)\|_2 + \|\mathbf{U}_1^\dagger\|_2\|(\widehat{\mathbf{U}}_y - \mathbf{U}_y)\|_2
\end{aligned}$$

where the last inequality follows from the fact that $\widehat{\mathbf{U}}_2, \widehat{\mathbf{U}}_2 - \mathbf{U}_2$ are submatrices of $\widehat{\mathbf{U}}_y$ and $\widehat{\mathbf{U}}_y - \mathbf{U}_y$, respectively. Therefore, we have $\|\widehat{\mathbf{U}}_2\| \le \|\widehat{\mathbf{U}}_y\|_2 = 1$, and $\|\widehat{\mathbf{U}}_2 - \mathbf{U}_2\|_2 \le \|\widehat{\mathbf{U}}_y - \mathbf{U}_y\|_2$. By the assumption in this lemma, we have,

$$\|\widehat{\mathbf{U}}_1 - \mathbf{U}_1\|_2 \le \|\widehat{\mathbf{U}}_y - \mathbf{U}_y\|_2 \le \frac{\sigma_S(\mathbf{U}_1)}{2} \tag{54}$$

We use a result from [5, Theorem 3.2] which states that a matrix $\mathbf{F}$ with rank S, and its perturbed matrix $\widetilde{\mathbf{F}} = \mathbf{F} + \mathbf{E}$ satisfy the following inequality:

$$\|\mathbf{F}^\dagger - \widetilde{\mathbf{F}}^\dagger\|_2 \le \frac{3\|\mathbf{E}\|_2}{\sigma_S(\mathbf{F})(\sigma_S(\mathbf{F}) - \|\mathbf{E}\|_2)}$$

provided the perturbation satisfies $\|\mathbf{E}\|_2 < \sigma_S(\mathbf{F})$. We use this result by substituting $\mathbf{F}$ with $\mathbf{U}$, and $\widetilde{\mathbf{F}}$ with $\widehat{\mathbf{U}}_1$.

From (54), the perturbation condition is satisfied and this result leads to:

$$\begin{aligned}
\|\widehat{\mathbf{U}}_1^\dagger - \mathbf{U}_1^\dagger\|_2 &\le \frac{3\|\widehat{\mathbf{U}}_1 - \mathbf{U}_1\|_2}{\sigma_S(\mathbf{U}_1)(\sigma_S(\mathbf{U}_1) - \|\widehat{\mathbf{U}}_1 - \mathbf{U}_1\|_2)} \\
&\le \frac{6\|\widehat{\mathbf{U}}_y - \mathbf{U}_y\|_2}{\sigma_S(\mathbf{U}_1)^2}
\end{aligned} \tag{55}$$

Therefore, we have that

$$\begin{aligned}
\|\mathbf{\Psi} - \widehat{\mathbf{\Psi}}\|_2 &\le \left(\frac{6}{\sigma_S(\mathbf{U}_1)^2} + \frac{1}{\sigma_S(\mathbf{U}_1)}\right)\|\widehat{\mathbf{U}}_y - \mathbf{U}_y\|_2 \\
&\le \frac{7\|\widehat{\mathbf{U}}_y - \mathbf{U}_y\|_2}{\sigma_S(\mathbf{U}_1)^2}
\end{aligned} \tag{56}$$

$\square$

Lemma 6. *Consider the measurement model in (14) such that (17) holds. Then the following bound is satisfied:*

$$\|\mathbf{\Psi} - \widehat{\mathbf{\Psi}}\|_2 \le \frac{14\sqrt{2S}\|\widetilde{\mathbf{N}}\|_2}{\sigma_S(\mathbf{B})\sigma_S(\mathbf{X})\sigma_S(\mathbf{U}_1)^2} \tag{57}$$

Proof. From (17) and (49), we have

$$\|\widetilde{\mathbf{N}}\|_2 \le \frac{\sigma_S(\mathbf{B})\sigma_S(\mathbf{X})\sigma_S(\mathbf{U}_1)}{8\sqrt{2S}} \le \frac{\sigma_S(\mathbf{B})\sigma_S(\mathbf{X})}{2} \tag{58}$$

where the second inequality follows from the fact that $\sigma_S(\mathbf{U}_1) \le 1$ and $S \ge 1$. By applying Lemma 4, (50) holds. Now, (50) and (58) together imply that $\|\mathbf{U}_y - \widehat{\mathbf{U}}_y\|_2 \le \sigma_S(\mathbf{U}_1)/2$. This ensures that the conditions of Lemma 5 are satisfied. Combining (53) and (50) leads to the desired result. $\square$

Lemma 7.

$$md(\mathcal{F}, \widehat{\mathcal{F}}) \le \frac{1}{2} md(\mathbf{\Psi}, \widehat{\mathbf{\Psi}}) \tag{59}$$

Proof. The proof follows directly from eq. (III.1) in [1] $\square$

Lemma 8. *Consider the measurement model in (14). If rank*$(\mathbf{BX}) = S$*, then*

$$md(\mathcal{F}, \widehat{\mathcal{F}}) \le \frac{S\|\mathbf{B}\|_2}{\sigma_S(\mathbf{B})}\|\mathbf{\Psi} - \widehat{\mathbf{\Psi}}\|_2 \tag{60}$$

Proof. Based on (9), $\mathbf{\Psi}$ is diagonalizable by the invertible matrix $\mathbf{P}$. Using Bauer-Fike theorem, [6], [4, Theorem 3.3] and Lemma 7, we have

$$md(\mathcal{F}, \widehat{\mathcal{F}}) \le \frac{1}{2}(2S-1)\kappa(\mathbf{P}^{-1})\|\mathbf{\Psi} - \widehat{\mathbf{\Psi}}\|_2 \tag{61}$$

where $\kappa(\mathbf{P}^{-1}) = \|\mathbf{P}\|_2\|\mathbf{P}^{-1}\|_2$. To bound $\kappa(\mathbf{P}^{-1})$, we use the fact that $\mathbf{U}_y = \mathbf{BP}$ and $\|\mathbf{U}_y\|_2 = 1$. This implies that

$$\kappa(\mathbf{P}^{-1}) \le \kappa(\mathbf{B}) = \frac{\|\mathbf{B}\|_2}{\sigma_S(\mathbf{B})} \tag{62}$$

$\square$

References

[1] W. Li, W. Liao, and A. Fannjiang, "Super-resolution limit of the esprit algorithm," *IEEE Transactions on Information Theory*, vol. 66, no. 7, pp. 4593–4608, 2020.

[2] P.-A. Wedin, "Perturbation bounds in connection with singular value decomposition," *BIT Numerical Mathematics*, vol. 12, no. 1, pp. 99–111, 1972.

[3] J. N. Franklin, *Matrix theory*. Courier Corporation, 2012.

[4] G. W. Stewart, "Matrix perturbation theory," 1990.

[5] P. C. Hansen, "The truncatedsvd as a method for regularization," *BIT Numerical Mathematics*, vol. 27, no. 4, pp. 534–553, 1987.

[6] F. L. Bauer and C. T. Fike, "Norms and exclusion theorems," *Numerische Mathematik*, vol. 2, no. 1, pp. 137–141, 1960.