SUPPLEMENTARY MATERIALS: TIGHTNESS OF SDP AND BURER-MONTEIRO FACTORIZATION FOR PHASE SYNCHRONIZATION IN HIGH-NOISE REGIME

ANDERSON YE ZHANG*

SM1. Proofs of Lemmas in Section 2. We defer the proof of Lemma 2.1 to Section SM5 as the lemma is a direct generalization of Lemma 12 of [SM2] and our proof follows theirs.

Proof of Lemma 2.2. To prove (2.8), let $\theta \in [0,\pi]$ be the angle between x and y. By the cosine formula of triangles, we have $\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos(\theta)$ and $\|x/\|x\| - y/\|y\|\|^2 = 2 - 2\cos(\theta)$. Consider the following scenarios.

- If $\|x\|, \|y\| \geq t$, since $\|x\|^2 + \|y\|^2 \geq 2\|x\|\|y\|$, we have

$$\|x-y\|^2 \geq 2\|x\|\|y\|(1-\cos(\theta)) \geq 2t^2(1-\cos(\theta)) = t^2\|x/\|x\| - y/\|y\|\|^2.$$

  Hence, $\|x/\|x\| - y/\|y\|\| \leq \|x-y\|/t$.
- If $\|y\| \geq t > \|x\|$ and $\cos(\theta) \geq 0$, define a function $f(a,b) = a^2+b^2-2ab\cos(\theta)$ for $a,b \in \mathbb{R}$. Note that for any $1 \geq a > 0, b \geq 1$, we have $f(a,b) \geq 1-\cos^2(\theta)$. This is because $f(a,b) \geq \min_{b'\geq 1} f(a,b') = f(a,1) = a^2+1-2a\cos(\theta) \geq \min_{1\geq a'>0} f(a',1) = f(\cos(\theta),1) = 1-\cos^2(\theta)$. Hence,

$$\begin{aligned}\frac{2\left\|x-y\right\|^2}{t^2} &= 2\left(\left(\frac{\|x\|}{t}\right)^2 + \left(\frac{\|y\|}{t}\right)^2 - \frac{\|x\|}{t}\frac{\|y\|}{t}\cos(\theta)\right)\\ &\geq 2(1-\cos^2(\theta))\\ &\geq 2(1-\cos(\theta))\\ &= \left\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\right\|^2.\end{aligned}$$

  Hence, $\|x/\|x\| - y/\|y\|\| \leq \sqrt{2}\left\|x-y\right\|/t$.
- If $\|y\| \geq t > \|x\|$ and $\cos(\theta) < 0$, we have $\left\|x-y\right\|^2 \geq \left\|y\right\|^2 \geq t^2$ and $\|x/\|x\| - y/\|y\|\| \leq 2$. Hence, $\|x/\|x\| - y/\|y\|\| \leq 2\left\|x-y\right\|/t$.
- If $\|y\| < t$, we have $\|x/\|x\| - y/\|y\|\| \leq 2 = 2\mathbb{I}\left\{\|y\| < t\right\}$.

The proof of (2.8) is complete.

To prove (2.9), we only need to consider scenarios $x = 0$ or $y = 0$, as otherwise (2.9) is reduced to (2.8). If $y = 0$, we have

$$\begin{aligned}&\left\|\left(\frac{x}{\|x\|}\mathbb{I}\left\{x\neq 0\right\} + u\mathbb{I}\left\{x=0\right\}\right) - \left(\frac{y}{\|y\|}\mathbb{I}\left\{y\neq 0\right\} + v\mathbb{I}\left\{y=0\right\}\right)\right\|\\ &= \left\|\left(\frac{x}{\|x\|}\mathbb{I}\left\{x\neq 0\right\} + u\mathbb{I}\left\{x=0\right\}\right) - v\right\| \leq 2 = 2\mathbb{I}\left\{\|y\| < t\right\}.\end{aligned}$$

If $x = 0$ and $y \neq 0$, we have

$$\left\|\left(\frac{x}{\|x\|}\mathbb{I}\left\{x\neq 0\right\} + u\mathbb{I}\left\{x=0\right\}\right) - \left(\frac{y}{\|y\|}\mathbb{I}\left\{y\neq 0\right\} + v\mathbb{I}\left\{y=0\right\}\right)\right\|$$

*Department of Statistics and Data Science, University of Pennsylvania (ayz@wharton.upenn.edu, https://andersonyezhang.github.io/).

$$= \left\| u - \frac{y}{\|y\|} \right\| \leq 2 = 2\mathbb{I}\{\|y\| \geq t\} + 2\mathbb{I}\{\|y\| < t\} = 2\mathbb{I}\{\|x-y\| \geq t\} + 2\mathbb{I}\{\|y\| < t\}$$
$$\leq \frac{2\|x-y\|}{t} + 2\mathbb{I}\{\|y\| < t\}.$$

The proof of (2.9) is complete. ∎

Proof of Lemma 2.5. Consider any $m \in \mathbb{N} \setminus \{1\}$. For simplicity, we write $\widehat{Z}^{\mathrm{BM},m}$ as $\widehat{Z}$ so that $\widehat{Z} = (\widehat{V}^{\mathrm{BM},m})^{\mathrm{H}} \widehat{V}^{\mathrm{BM},m}$.

First, we are going to show

$$\ell(\widehat{V}^{\mathrm{BM},m}, z^*) \leq \frac{4}{n^2} \operatorname{Tr}(z^* z^{*\mathrm{H}}(z^* z^{*\mathrm{H}} - \widehat{Z})). \tag{SM1.1}$$

Define $b = n^{-1} \sum_{j=1}^n \widehat{V}_j^{\mathrm{BM},m} z_j^* = n^{-1} \widehat{V}^{\mathrm{BM},m} z^* \in \mathbb{C}^m$. If $b = 0$, we have

$$\begin{aligned}
\operatorname{Tr}(z^* z^{*\mathrm{H}}(z^* z^{*\mathrm{H}} - \widehat{Z})) &= \operatorname{Tr}(z^*(z^*)^{\mathrm{H}} z^* (z^*)^{\mathrm{H}}) - \operatorname{Tr}(z^* z^{*\mathrm{H}} (\widehat{V}^{\mathrm{BM},m})^{\mathrm{H}} \widehat{V}^{\mathrm{BM},m}) \\
&= n \operatorname{Tr}(z^* z^{*\mathrm{H}}) - \operatorname{Tr}(z^* (nb)^{\mathrm{H}} \widehat{V}^{\mathrm{BM},m}) \\
&= n^2.
\end{aligned}$$

Note that $\ell(\widehat{V}^{\mathrm{BM},m}, z^*) \leq n^{-1} \sum_{j \in [n]} 4 = 4$. Then (SM1.1) holds. In the following, we assume $b \neq 0$. From Lemma 2.2, we have for any $x, y \in \mathbb{C}^m$ such that $x \neq 0$ and $\|y\| = 1$, $\|x / \|x\| - y\| \leq 2\|x - y\|$. Hence, we have

$$\begin{aligned}
\ell(\widehat{V}^{\mathrm{BM},m}, z^*) &= \min_{a \in \mathbb{C}^n : \|a\|^2 = 1} \frac{1}{n} \sum_{j=1}^n \|\widehat{V}_j^{\mathrm{BM},m} z_j^* - a\|^2 \\
&= \min_{a \in \mathbb{C}^n \setminus \{0\}} \frac{1}{n} \sum_{j=1}^n \|\widehat{V}_j^{\mathrm{BM},m} z_j^* - a/\|a\|\|^2 \\
&\leq \min_{a \in \mathbb{C}^n \setminus \{0\}} \frac{4}{n} \sum_{j=1}^n \|\widehat{V}_j^{\mathrm{BM},m} z_j^* - a\|^2.
\end{aligned}$$

Since the minimum of the above display is achieved when a is the arithmetic mean of $\{\widehat{V}_j^{\mathrm{BM},m} z_j^*\}_{j \in [n]}$, i.e., b, we have

$$\begin{aligned}
\ell(\widehat{V}^{\mathrm{BM},m}, z^*) &\leq \frac{4}{n} \sum_{j=1}^n \|\widehat{V}_j^{\mathrm{BM},m} z_j^* - b\|^2 \\
&= \frac{2}{n^2} \sum_{j=1}^n \sum_{l=1}^n \left(\|\widehat{V}_j^{\mathrm{BM},m} z_j^* - b\|^2 + \|\widehat{V}_l^{\mathrm{BM},m} z_l^* - b\|^2 \right) \\
&= \frac{2}{n^2} \sum_{j=1}^n \sum_{l=1}^n \|\widehat{V}_j^{\mathrm{BM},m} z_j^* - \widehat{V}_l^{\mathrm{BM},m} z_l^*\|^2 \\
&= \frac{4}{n^2} \sum_{j=1}^n \sum_{l=1}^n (1 - \overline{z}_j^* z_l^* (\widehat{V}_j^{\mathrm{BM},m})^{\mathrm{H}} \widehat{V}_l^{\mathrm{BM},m}) \\
&= \frac{4}{n^2} \operatorname{Tr}(z^* z^{*\mathrm{H}}(z^* z^{*\mathrm{H}} - \widehat{Z})).
\end{aligned}$$

Therefore, (SM1.1) holds.

Now it remains to upper bound $\mathrm{Tr}(z^* z^{*\mathrm{H}}(z^* z^{*\mathrm{H}} - \widehat{Z}))$. By the definition (1.6), we have $\mathrm{Tr}(Y\widehat{Z}) \geq \mathrm{Tr}(Y z^* z^{*\mathrm{H}})$. Rearranging this inequality, we obtain $\mathrm{Tr}(Y(\widehat{Z} - z^* z^{*\mathrm{H}})) \geq 0$. With (1.2), we have

$$
\begin{aligned}
\mathrm{Tr}(z^* z^{*\mathrm{H}}(z^* z^{*\mathrm{H}} - \widehat{Z})) &\leq \mathrm{Tr}\left((Y - z^* z^{*\mathrm{H}})(\widehat{Z} - z^* z^{*\mathrm{H}})\right) \\
&= \sigma \,\mathrm{Tr}\left(W(\widehat{Z} - z^* z^{*\mathrm{H}})\right) \\
&\leq \sigma \left|\mathrm{Tr}\left(W\widehat{Z}\right)\right| + \sigma \left|\mathrm{Tr}\left(W z^* z^{*\mathrm{H}}\right)\right| \\
&\leq \sigma \left\|W\right\| \mathrm{Tr}\left(\widehat{Z}\right) + \sigma \left\|W\right\| \mathrm{Tr}\left(z^* z^{*\mathrm{H}}\right) \\
&= 2n\sigma \left\|W\right\|.
\end{aligned}
$$

Here, the last inequality is due to the following facts. For any two matrices $A, B \in \mathbb{C}^{n\times n}$, $\mathrm{Tr}(AB) \leq \|A\| \|B\|_*$, where $\|B\|_*$ is the nuclear norm of B that is equal to the summation of all its singular values. If B is further assumed to be positive semi-definite, we have $\|B\|_* = \mathrm{Tr}(B)$. In our setting, $\widehat{Z}$ is positive semi-definite as $\min_{u\in\mathbb{C}^n} u^H \widehat{Z} u = \min_{u\in\mathbb{C}^n} u^{\mathrm{H}} (\widehat{V}^{\mathrm{BM},m})^{\mathrm{H}} \widehat{V}^{\mathrm{BM},m} u \geq 0$, and so is $z^*(z^*)^{\mathrm{H}}$.

Consequently, we have $\ell(\widehat{V}^{\mathrm{BM},m}, z^*) \leq \frac{8\sigma\|W\|}{n}$. The upper bound for $\ell_1(\widehat{z}^{\mathrm{MLE}}, z^*)$ can be established following the same steps as above and hence its proof is omitted. ∎

SM2. Proofs of Lemmas in Section 3.2.

Proof of Lemma 3.1. Consider the following scenarios. If $|x|, |y| \leq t$, we have $|g_t(x) - g_t(y)| = \frac{|x-y|}{t}$ by definition. If $|x|, |y| \geq t$, then

$$
|g_t(x) - g_t(y)| = \left|\frac{x}{|x|} - \frac{y}{|y|}\right|,
$$

Let $\theta \in [0, \pi]$ be the angle between x and y on the complex plane. By the cosine formula of triangles, we have $|x-y|^2 = |x|^2 + |y|^2 - 2|x||y|\cos(\theta)$ and $|g_t(x) - g_t(y)|^2 = 2 - 2\cos(\theta)$. Since $|x|^2 + |y|^2 \geq 2|x||y|$, we have

$$
|x - y|^2 \geq 2|x||y|(1 - \cos(\theta)) \geq 2t^2(1 - \cos(\theta)) = t^2 \left|g_t(x) - g_t(y)\right|^2,
$$

which yields the desired result. If $|x| \geq t > |y|$, then

$$
|g_t(x) - g_t(y)| = \left|\frac{x}{|x|} - \frac{y}{t}\right|.
$$

By using the cosine formula again, we have $|g_t(x) - g_t(y)|^2 = 1 + \frac{|y|^2}{t^2} - 2\frac{|y|}{t}\cos(\theta)$ and $\left|\frac{x}{t} - \frac{y}{t}\right|^2 = \frac{|x|^2}{t^2} + \frac{|y|^2}{t^2} - 2\frac{|x||y|}{t^2}\cos(\theta)$. Then,

$$
\begin{aligned}
\frac{|x - y|^2}{t^2} - |g_t(x) - g_t(y)|^2 &= \left|\frac{x}{t} - \frac{y}{t}\right|^2 - |g_t(x) - g_t(y)|^2 \\
&= \frac{|x|^2}{t^2} - 1 - 2\frac{|x||y|}{t^2}\cos(\theta) + 2\frac{|y|}{t}\cos(\theta) \\
&= \left(\frac{|x|}{t} - 1\right)\left(\frac{|x|}{t} + 1\right) - 2\left(\frac{|x|}{t} - 1\right)\frac{|y|}{t}\cos(\theta) \\
&= \left(\frac{|x|}{t} - 1\right)\left(\frac{|x|}{t} + 1 - 2\frac{|y|}{t}\cos(\theta)\right)
\end{aligned}
$$

$$\geq 0,$$

where the last inequality is due to that $\frac{|x|}{t} \geq 1 > \frac{|y|}{t} \geq 0$ and $\cos(\theta) \leq 1$. The scenario $|y| \geq t > |x|$ can be proved similarly. ◻

Proof of Lemma 3.2. We prove the properties sequentially.

1. Recall the definition of G in (3.9). For any $j \in [n]$, by Lemma 3.1, we have

$$\begin{aligned} \left|[G(x,s,t)]_j - [G(y,s,t)]_j\right| &= \left|g_t(z_j^* s + \sigma[Wx]_j) - g_t(z_j^* s + \sigma[Wy]_j)\right| \\ &\leq t^{-1}\left|\left(z_j^* s + \sigma[Wx]_j\right) - \left(z_j^* s + \sigma[Wx]_j\right)\right| \\ &= t^{-1}\sigma\left|[W(x-y)]_j\right|. \end{aligned}$$

   Summing over all $j \in [n]$, we have

$$\begin{aligned} \left\|G(x,s,t) - G(y,s,t)\right\|^2 &\leq \sum_{j\in[n]} \left|[G(x,s,t)]_j - [G(y,s,t)]_j\right|^2 \\ &\leq t^{-2}\sigma^2 \sum_{j\in[n]} \left|[W(x-y)]_j\right|^2 \\ &= t^{-2}\sigma^2 \left\|W(x-y)\right\|^2 \\ &\leq t^{-2}\sigma^2 \left\|W\right\|^2 \left\|x-y\right\|^2. \end{aligned}$$

2. Using the first property, for any $T \in \mathbb{N}$, we have

$$\begin{aligned} \left\|z^{(T+1)} - z^{(T)}\right\| &= \left\|G(z^{(T)},s,t) - G(z^{(T-1)},s,t)\right\| \\ &\leq t^{-1}\sigma\left\|W\right\|\left\|z^{(T)} - z^{(T-1)}\right\| \\ &\leq \frac{1}{2}\left\|z^{(T)} - z^{(T-1)}\right\|, \end{aligned}$$

   where the last inequality is due to the assumption $t \geq 2\sigma\left\|W\right\|$.

3. Consider the sequence $z^{(0)} = z^*$ and $z^{(T)} = G(z^{(T-1)},s,t)$ for all $T \in \mathbb{N}$. By the second property, we have $\left\|z^{(T+1)} - z^{(T)}\right\| \leq \frac{1}{2}\left\|z^{(T)} - z^{(T-1)}\right\|$ for all $T \in \mathbb{N}$. Note that $\{z^{(T)}\}$ is a sequence in $\mathbb{C}^n_{\leq 1}$, a complete metric space under $\left\|\cdot\right\|$. Hence, the sequence converges to a limit $z^{(\infty)} \in \mathbb{C}^n_{\leq 1}$ which satisfies $z^{(\infty)} = G(z^{(\infty)},s,t)$. Hence, $z^{(\infty)}$ is a fixed point of $G(\cdot,s,t)$. Now we have proved the existence of the fixed point. To prove the uniqueness, note that if there exists another $z' \in \mathbb{C}^n_{\leq 1}$ such that $z' = G(z',s,t)$, we have

$$\left\|z^{(\infty)} - z'\right\| = \left\|G(z^{(\infty)},s,t) - G(z',s,t)\right\| \leq t^{-1}\sigma\left\|z^{(\infty)} - z'\right\| \leq \left\|z^{(\infty)} \text{[redacted]} z'\right\|/2,$$

   by the first property. Hence, $\left\|z^{(\infty)} - z'\right\| = 0$ which means $z^{(\infty)} = z'$.

4. For any $j \in [n]$, we have

$$\begin{aligned} \left|[z^* s + \sigma W z]_j - [z^* s' + \sigma W z']_j\right| &\leq \left|z_j^* s - z_j^* s'\right| + \sigma\left|[W(z-z')]_j\right| \\ &\leq \left|s - s'\right| + \sigma\left|[W(z-z')]_j\right|. \end{aligned}$$

   Summing over all $j \in [n]$, we have

$$\left\|(z^* s + \sigma W z) - (z^* s' + \sigma W z')\right\|^2 \leq \sum_{j\in[n]} \left(\left|s-s'\right| + \sigma\left|[W(z-z')]_j\right|\right)^2$$

$$
\begin{aligned}
&\le \sum_{j\in[n]} \left(2\left|s-s'\right|^2 + 2\sigma^2 \left|[W(z-z')]_j\right|^2\right)\\
&\le 2n\left|s-s'\right|^2 + 2\sigma^2 \left\|W\right\|^2 \left\|z-z'\right\|^2 . \qquad \text{(SM2.1)}
\end{aligned}
$$

Note that for any $j\in[n]$, we have $z_j = [G(z,s,t)]_j = g_t([z^*s+\sigma Wz]_j)$ and similarly $z'_j = g_t([z^*s'+\sigma Wz']_j)$. Hence, by Lemma 3.1, we have

$$
\left|z_j - z'_j\right| \le t^{-1}\left|[z^*s+\sigma Wz]_j - [z^*s'+\sigma Wz']_j\right| .
$$

Summing over all $j\in[n]$, by (SM2.1), we have

$$
\begin{aligned}
\left\|z-z'\right\|^2 &\le t^{-2}\left\|(z^*s+\sigma Wz)-(z^*s'+\sigma Wz')\right\|^2\\
&\le 2nt^{-2}\left|s-s'\right|^2 + 2\sigma^2 t^{-2}\left\|W\right\|^2\left\|z-z'\right\|^2\\
&\le 2nt^{-2}\left|s-s'\right|^2 + \frac{1}{2}\left\|z-z'\right\|^2 ,
\end{aligned}
$$

where the last inequality is due to the assumption $t\ge 2\sigma\left\|W\right\|$. After rearrangement, we have $\left\|z-z'\right\|^2 \le 4nt^{-2}\left|s-s'\right|^2$. From (SM2.1), we have

$$
\begin{aligned}
\left\|(z^*s+\sigma Wz)-(z^*s'+\sigma Wz')\right\|^2 &\le 2n\left|s-s'\right|^2 + 2\sigma^2\left\|W\right\|^2\left(4nt^{-2}\left|s-s'\right|^2\right)\\
&\le 4n\left|s-s'\right|^2 , \qquad \square
\end{aligned}
$$

where the last inequality is by $t\ge 2\sigma\left\|W\right\|$.

Proof of Lemma 3.3. Consider any $j\in[n]$. If $|z_j^*s+\sigma[Wz]_j|\ge t$, we have $[G(z,s,t)]_j = g_t(z_j^*s+\sigma[Wz]_j) = (z_j^*s+\sigma[Wz]_j)/|z_j^*s+\sigma[Wz]_j| = [F_1'(z,s)]_j$. If $|z_j^*s+\sigma[Wz]_j|\ge t$ is not satisfied, we have $|[F_1'(z,s)]_j| = 1$ and $|[G(z,s,t)]_j|\le 1$. Hence,

$$
\begin{aligned}
|[F_1'(z,s)]_j - [G(z,s,t)]_j| &= |[F_1'(z,s)]_j - [G(z,s,t)]_j|\,\mathbb{I}\left\{|z_j^*s+\sigma[Wz]_j|<t\right\}\\
&\le 2\mathbb{I}\left\{|z_j^*s+\sigma[Wz]_j|<t\right\}.
\end{aligned}
$$

Summing over all $j\in[n]$, we have

$$
\left\|F_1'(z,s)-G(z,s,t)\right\|^2 = \sum_{j\in[n]}\left|[F_1'(z,s)]_j - [G(z,s,t)]_j\right|^2 \le 4\sum_{j\in[n]}\mathbb{I}\left\{|z_j^*s+\sigma[Wz]_j|<t\right\}. \qquad \square
$$

Proof of Lemma 3.4. Consider any $t>0$. From (3.7), we have $\widehat{z}^{\mathrm{MLE}} = F_1'(\widehat{z}^{\mathrm{MLE}},\widehat{s})$. Then

$$
\begin{aligned}
\left\|\widehat{z}^{\mathrm{MLE}} - z\right\| &= \left\|F_1'(\widehat{z}^{\mathrm{MLE}},\widehat{s}) - G(z,\widehat{s},t)\right\|\\
&\le \left\|F_1'(\widehat{z}^{\mathrm{MLE}},\widehat{s}) - F_1'(z,\widehat{s})\right\| + \left\|F_1'(z,\widehat{s}) - G(z,\widehat{s},t)\right\|\\
&\le \left\|F_1'(\widehat{z}^{\mathrm{MLE}},\widehat{s}) - F_1'(z,\widehat{s})\right\| + \sqrt{4\sum_{j\in[n]}\mathbb{I}\left\{|z_j^*\widehat{s}+\sigma[Wz]_j|<t\right\}},
\end{aligned}
$$

where the last inequality is due to Lemma 3.3. Hence,

$$
\left\|\widehat{z}^{\mathrm{MLE}} - z\right\|^2 \le 2\left\|F_1'(\widehat{z}^{\mathrm{MLE}},\widehat{s}) - F_1'(z,\widehat{s})\right\|^2 + 8\sum_{j\in[n]}\mathbb{I}\left\{|z_j^*\widehat{s}+\sigma[Wz]_j|<t\right\}.
$$

Recall the definition of F_1' in (3.5). For any $j \in [n]$, note that $[F_1'(\widehat{z}^{\mathrm{MLE}}, \widehat{s})]_j$ and $[F_1'(z, \widehat{s})]_j$ can be written as

$$[F_1'(\widehat{z}^{\mathrm{MLE}}, \widehat{s})]_j = \frac{z_j^*\widehat{s} + \sigma[W\widehat{z}^{\mathrm{MLE}}]_j}{|z_j^*\widehat{s} + \sigma[W\widehat{z}^{\mathrm{MLE}}]_j|}\mathbb{I}\left\{z_j^*\widehat{s} + \sigma[W\widehat{z}^{\mathrm{MLE}}]_j \neq 0\right\} + \widehat{z}_j^{\mathrm{MLE}}\mathbb{I}\left\{z_j^*\widehat{s} + \sigma[W\widehat{z}^{\mathrm{MLE}}]_j = 0\right\},$$

$$[F_1'(z, \widehat{s})]_j = \frac{z_j^*\widehat{s} + \sigma[Wz]_j}{|z_j^*\widehat{s} + \sigma[Wz]_j|}\mathbb{I}\left\{z_j^*\widehat{s} + \sigma[Wz]_j \neq 0\right\} + z_j\mathbb{I}\left\{z_j^*\widehat{s} + \sigma[Wz]_j = 0\right\}.$$

By applying (2.9) of Lemma 2.2,

$$\begin{aligned}\left|[F_1'(\widehat{z}^{\mathrm{MLE}}, \widehat{s})]_j - [F_1'(z, \widehat{s})]_j\right| &\leq \frac{2\left|\left(z_j^*\widehat{s} + \sigma[W\widehat{z}^{\mathrm{MLE}}]_j\right) - \left(z_j^*\widehat{s} + \sigma[Wz]_j\right)\right|}{t} + 2\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\} \\ &\leq \frac{2\sigma\left|[W(\widehat{z}^{\mathrm{MLE}} - z)]_j\right|}{t} + 2\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\}.\end{aligned}$$

Summing over all $j \in [n]$, we have

$$\begin{aligned}\left\|F_1'(\widehat{z}^{\mathrm{MLE}}, \widehat{s}) - F_1'(z, \widehat{s})\right\|^2 &= \sum_{j\in[n]}\left|[F_1'(\widehat{z}^{\mathrm{MLE}}, \widehat{s})]_j - [F_1'(z, \widehat{s})]_j\right|^2 \\ &= \sum_{j\in[n]}\left(\frac{4\sigma^2\left|[W(\widehat{z}^{\mathrm{MLE}} - z)]_j\right|^2}{t^2} + 4\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\}\right) \\ &= \frac{4\sigma^2}{t^2}\left\|W(\widehat{z}^{\mathrm{MLE}} - z)\right\|^2 + 4\sum_{j\in[n]}\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\} \\ &\leq \frac{4\sigma^2}{t^2}\left\|W\right\|^2\left\|\widehat{z}^{\mathrm{MLE}} - z\right\|^2 + 4\sum_{j\in[n]}\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\}.\end{aligned}$$

Hence,

$$\begin{aligned}\left\|\widehat{z}^{\mathrm{MLE}} - z\right\|^2 &\leq 2\left(\frac{4\sigma^2}{t^2}\left\|W\right\|^2\left\|\widehat{z}^{\mathrm{MLE}} - z\right\|^2 + 4\sum_{j\in[n]}\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\}\right) \\ &\quad + 8\sum_{j\in[n]}\mathbb{I}\left\{|z_j^*\widehat{s} + \sigma[Wz]_j| < t\right\} \\ &= \frac{8\sigma^2}{t^2}\left\|W\right\|^2\left\|\widehat{z}^{\mathrm{MLE}} - z\right\|^2 + 16\sum_{j\in[n]}\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\}.\end{aligned}$$

When $t \geq 4\sigma\left\|W\right\|$, we have $\frac{8\sigma^2}{t^2}\left\|W\right\|^2 \leq 1/2$ and the above display leads to $\left\|\widehat{z}^{\mathrm{MLE}} - z\right\|^2 \leq 32\sum_{j\in[n]}\mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma[Wz]_j\right| < t\right\}$. ∎

Proof of Lemma 3.6. Recall the definitions of G in (3.9) and g_t in (3.8). Note that for any $t > 0, a \in \mathbb{C}_1, x \in \mathbb{C}$, we have $ag_t(x) = g_t(ax)$. Hence, for any $z \in \mathbb{C}^n_{\leq 1}, s \in \mathbb{C}, t > 0$, $a \in \mathbb{C}_1$, and $j \in [n]$, we have $a[G(z, s, t)]_j = ag_t([z^*s + \sigma Wz]_j) = g_t(a[z^*s + \sigma Wz]_j) = g_t([z^*(as) + \sigma W(az)]_j)$. As a result,

$$\text{if } z = G(z, s, t), \text{ then } az = G(az, as, t).$$

This means that a fixed point of $G(\cdot, s, t)$ is also a fixed point of $G(\cdot, as, t)$.

Recall the definition of $\widehat{s}$ in (3.6). We only need to study the case that $\widehat{s} \neq 0$ as otherwise $G(\cdot, |\widehat{s}|, \cdot) = G(\cdot, \widehat{s}, \cdot)$ and Lemma 3.6 is identical to Lemma 3.5. Since $\widehat{s} \neq 0$, $\widehat{s}/|\widehat{s}| \in \mathbb{C}_1$ is well-defined. For any $\delta \geq \frac{2\sigma\|W\|}{n}$, let $z \in \mathbb{C}^n_{\leq 1}$ be the fixed point of $G(\cdot, |\widehat{s}|, 2\delta n)$. Then we have $\frac{\widehat{s}}{|\widehat{s}|} z \in \mathbb{C}^n_{\leq 1}$ and

$$\frac{\widehat{s}}{|\widehat{s}|} z = G\left(\frac{\widehat{s}}{|\widehat{s}|} z, \frac{\widehat{s}}{|\widehat{s}|}|\widehat{s}|, 2\delta n\right) = G\left(\frac{\widehat{s}}{|\widehat{s}|} z, \widehat{s}, 2\delta n\right).$$

That is, $\frac{\widehat{s}}{|\widehat{s}|} z$ is the fixed point of $G(\cdot, \widehat{s}, 2\delta n)$. By Lemma 3.5, we have

$$\begin{aligned}
\frac{1}{n} \sum_{j\in[n]} \mathbb{I}\left\{\left|[Y\widehat{z}^{\mathrm{MLE}}]_j\right| < \delta n\right\} &\leq \frac{9}{n} \sum_{j\in[n]} \mathbb{I}\left\{\left|z_j^*\widehat{s} + \sigma\left[W\frac{\widehat{s}}{|\widehat{s}|} z\right]_j\right| < 2\delta n\right\} \\
&= \frac{9}{n} \sum_{j\in[n]} \mathbb{I}\left\{\left|\frac{\widehat{s}}{|\widehat{s}|}\left(z_j^*|\widehat{s}| + \sigma\left[Wz\right]_j\right)\right| < 2\delta n\right\} \\
&= \frac{9}{n} \sum_{j\in[n]} \mathbb{I}\left\{\left|z_j^*|\widehat{s}| + \sigma\left[Wz\right]_j\right| < 2\delta n\right\}. \qquad \square
\end{aligned}$$

SM3. Proofs of Lemmas in Section 3.3.

The following lemma is a counterpart of Lemma 3.2 but for $G^{(-j)}$ instead of G. Then Lemma 3.9 is the direct consequence of the third properties of Lemmas 3.2 and SM3.1.

LEMMA SM3.1. *Consider any $j \in [n]$. The function $G^{(-j)}(\cdot,\cdot,\cdot)$ has the following properties:*

1. *For any $x, y \in \mathbb{C}^n$ and for any $s \in \mathbb{C}, t > 0$, we have*

$$\left\|G^{(-j)}(x,s,t) - G^{(-j)}(y,s,t)\right\| \leq t^{-1}\sigma\|W\|\|x-y\|^2.$$

2. *For any $s \in \mathbb{C}, t \geq 2\sigma\|W\|$, and for any $z^{(0,-j)} \in \mathbb{C}^n_{\leq 1}$, define $z^{(T,-j)} = G^{(-j)}(z^{(T-1,-j)}, s, t)$ for all $T \in \mathbb{N}$. Then*

$$\left\|z^{(T+1,-j)} - z^{(T,-j)}\right\| \leq \frac{1}{2}\left\|z^{(T,-j)} - z^{(T-1,-j)}\right\|, \forall T \in \mathbb{N}.$$

3. *For any $s \in \mathbb{C}, t \geq 2\sigma\|W\|$, $G(\cdot, s, t)$ has exactly one fixed point. That is, there exists one and only one $z \in \mathbb{C}^n_{\leq 1}$ such that $z = G^{(-j)}(z, s, t)$. In addition, z can be achieved by iteratively applying $G^{(-j)}(\cdot, s, t)$ starting from z^*. That is, let $z^{(0,-j)} = z^*$ and define $z^{(T,-j)} = G^{(-j)}(z^{(T-1,-j)}, s, t)$ for all $T \in \mathbb{N}$. We have $z = \lim_{T\to\infty} G^{(-j)}(z^{(T,-j)}, s, t)$.*

Proof. Note that $\left\|W^{(-j)}\right\| \leq \|W\|$ since $W^{(-j)}$ is obtained from W by zeroing out the jth row and column. With this, the lemma can be proved following the exact same argument as in the proof of Lemma 3.2, and hence is omitted here. $\square$

Proof of Lemma 3.10. Consider any $T \in \mathbb{N}$. For any $k \in [n]$, by Lemma 3.1, we have

$$\begin{aligned}
&\left|z_k^{(T)} - z_k^{(T,-j)}\right| \\
&= \left|[G(z^{(T-1)}, s, t)]_k - [G^{(-j)}(z^{(T-1,-j)}, s, t)]_k\right|
\end{aligned}$$

$$
\begin{aligned}
&= \left| g_t([z^* s + \sigma W z^{(T-1)}]_k) - g_t([z^* s + \sigma W^{(-j)} z^{(T-1,-j)}]_k) \right| \\
&\leq t^{-1} \left| [z^* s + \sigma W z^{(T-1)}]_k - [z^* s + \sigma W^{(-j)} z^{(T-1,-j)}]_k \right| \\
&= t^{-1}\sigma \left| [W z^{(T-1)}]_k - [W^{(-j)} z^{(T-1,-j)}]_k \right| \\
&= t^{-1}\sigma \left| [W z^{(T-1)}]_k - [W^{(-j)} z^{(T-1)}]_k + [W^{(-j)} z^{(T-1)}]_k - [W^{(-j)} z^{(T-1,-j)}]_k \right| \\
&= t^{-1}\sigma \left| [(W - W^{(-j)}) z^{(T-1)}]_k + [W^{(-j)} (z^{(T-1)} - z^{(T-1,-j)})]_k \right| \\
&\leq t^{-1}\sigma \left| [(W - W^{(-j)}) z^{(T-1)}]_k \right| + t^{-1}\sigma \left| [W^{(-j)} (z^{(T-1)} - z^{(T-1,-j)})]_k \right|.
\end{aligned}
$$

If $k \neq j$, we have $[(W - W^{(-j)}) z^{(T-1)}]_k = W_{kj} z_j^{(j)}$. Then the above display becomes

$$
\begin{aligned}
\left| z_k^{(T)} - z_k^{(T,-j)} \right| &\leq t^{-1}\sigma \left| W_{kj} z_j^{(j)} \right| + t^{-1}\sigma \left| [W^{(-j)} (z^{(T-1)} - z^{(T-1,-j)})]_k \right| \\
&\leq t^{-1}\sigma \left| W_{kj} \right| + t^{-1}\sigma \left| [W^{(-j)} (z^{(T-1)} - z^{(T-1,-j)})]_k \right|,
\end{aligned}
$$

where in the last inequality we use $|z_j^{(j)}| \leq 1$ as $z^{(j)} \in \mathbb{C}^n_{\leq 1}$. Summing over all $k \in [n]$ such that $k \neq j$, we have

$$
\begin{aligned}
&\sum_{k \in [n]: k \neq j} \left| z_k^{(T)} - z_k^{(T,-j)} \right|^2 \\
&\leq \sum_{k \in [n]: k \neq j} \left(2t^{-2}\sigma^2 \left| W_{kj} \right|^2 + 2t^{-2}\sigma^2 \left| [W^{(-j)} (z^{(T-1)} - z^{(T-1,-j)})]_k^2 \right| \right) \\
&\leq \sum_{k \in [n]} \left(2t^{-2}\sigma^2 \left| W_{kj} \right|^2 + 2t^{-2}\sigma^2 \left| [W^{(-j)} (z^{(T-1)} - z^{(T-1,-j)})]_k^2 \right| \right) \\
&= 2t^{-2}\sigma^2 \left\| W_j \right\|^2 + 2t^{-2}\sigma^2 \left\| W^{(-j)} (z^{(T-1)} - z^{(T-1,-j)} \right\|^2 \\
&\leq 2t^{-2}\sigma^2 \left\| W_j \right\|^2 + 2t^{-2}\sigma^2 \left\| W^{(-j)} \right\|^2 \left\| z^{(T-1)} - z^{(T-1,-j)} \right\|^2 \\
&\leq 2t^{-2}\sigma^2 \left\| W_j \right\|^2 + 2t^{-2}\sigma^2 \left\| W \right\|^2 \left\| z^{(T-1)} - z^{(T-1,-j)} \right\|^2,
\end{aligned}
$$

where in the last inequality, $\left\| W^{(-j)} \right\| \leq \|W\|$ due to that $W^{(-j)}$ is obtained from W by zeroing out its jth row and column. On the other hand, $\left| z_j^{(T)} - z_j^{(T,-j)} \right| \leq 2$. Hence,

$$
\begin{aligned}
\left\| z^{(T)} - z^{(T,-j)} \right\|^2 &\leq 4 + \sum_{k \in [n]: k \neq j} \left| z_k^{(T)} - z_k^{(T,-j)} \right|^2 \\
&\leq 4 + 2t^{-2}\sigma^2 \left\| W_j \right\|^2 + 2t^{-2}\sigma^2 \left\| W \right\|^2 \left\| z^{(T-1)} - z^{(T-1,-j)} \right\|^2 \\
&\leq 4 + 2t^{-2}\sigma^2 \left\| W \right\|^2 + 2t^{-2}\sigma^2 \left\| W \right\|^2 \left\| z^{(T-1)} - z^{(T-1,-j)} \right\|^2,
\end{aligned}
$$

where in the last inequality we use a fact that the operator norm of matrix is greater or equal to the norm of each column. When $t \geq 2\sigma \|W\|$, we have $2t^{-2}\sigma^2 \left\| W \right\|^2 \leq 1/2$ and

$$
\left\| z^{(T)} - z^{(T,-j)} \right\|^2 \leq \frac{9}{2} + \frac{1}{2} \left\| z^{(T-1)} - z^{(T-1,-j)} \right\|^2.
$$

Note that $\left\|z^{(0)} - z^{(0,-j)}\right\|^2 = 0$, by mathematical induction, it is easy to verify $\left\|z^{(T)} - z^{(T,-j)}\right\|^2 \leq 9, \forall T \in \mathbb{N}$. Let $T \to \infty$, we have $\left\|z - z^{(-j)}\right\|^2 \leq 9$. ∎

SM4. Proofs of Lemmas in Section 3.4.

Proof of Lemma 3.12. From Corollary 2.4, we have

$$\ell_m(\widehat{V}^{\mathrm{BM},m}, \widehat{z}^{\mathrm{MLE}}) \leq \frac{8}{n} \sum_{j\in[n]} \mathbb{I}\left\{\left|[Y\widehat{z}^{\mathrm{MLE}}]_j\right| < \delta n\right\}.$$

For each $k = 0, 1, 2, \ldots, \lceil n\epsilon/h \rceil$, let $z_{s_k} \in \mathbb{C}^n_{\leq 1}$ be the fixed point of $G(\cdot, s_k, 2\delta n)$. Then by Corollary 3.8, we have

$$\begin{aligned}
&\ell_m(\widehat{V}^{\mathrm{BM},m}, \widehat{z}^{\mathrm{MLE}}) \\
&\leq 72 \sum_{0\leq k\leq \lceil n\epsilon/h \rceil} \left(\frac{1}{n}\sum_{j\in[n]} \mathbb{I}\left\{\sigma\left|[Wz_{s_k}]_j\right| > s_k - 4\delta n\right\}\right) + \frac{72h^2}{\delta^2 n^2}\mathbb{I}\left\{h > \delta\sqrt{n}\right\}.
\end{aligned}$$

Since $2\delta n > 2\sigma\|W\|$, for each $k = 0, 1, 2, \ldots, \lceil n\epsilon/h \rceil$, Proposition 3.11 can be applied, leading to

$$\begin{aligned}
\frac{1}{n}\sum_{j\in[n]} \mathbb{I}\left\{\sigma\left|[Wz_{s_k}]_j\right| > s_k - 4\delta n\right\} &\leq \frac{1}{n}\sum_{j\in[n]} \mathbb{I}\left\{\sigma\left|W_{j\cdot}z_{s_k}^{(-j)}\right| > s_k - 4\delta n - 3\sigma\|W\|\right\} \\
&\leq \frac{1}{n}\sum_{j\in[n]} \mathbb{I}\left\{\sigma\left|W_{j\cdot}z_{s_k}^{(-j)}\right| > (1-\epsilon)n - h - 4\delta n - 3\sigma\|W\|\right\} \\
&= \frac{1}{n}\sum_{j\in[n]} \mathbb{I}\left\{\sigma\left|W_{j\cdot}z_{s_k}^{(-j)}\right| > \left(1 - \epsilon - 4\delta - \frac{3\sigma\|W\|}{n}\right)n - h\right\},
\end{aligned}$$

where in the last inequality, we use $\min_{0\leq k\leq\lceil n\epsilon/h\rceil} s_k \geq n - (n\epsilon/h + 1)h = (1-\epsilon)n - h$. Hence, we have

$$\begin{aligned}
&\ell_m(\widehat{V}^{\mathrm{BM},m}, \widehat{z}^{\mathrm{MLE}}) \\
&\leq 72 \sum_{0\leq k\leq \lceil n\epsilon/h \rceil} \left(\frac{1}{n}\sum_{j\in[n]} \mathbb{I}\left\{\sigma\left|W_{j\cdot}z_{s_k}^{(-j)}\right| > \left(1 - \epsilon - 4\delta - \frac{3\sigma\|W\|}{n}\right)n - h\right\}\right) + \frac{72h^2}{\delta^2 n^2}\mathbb{I}\left\{h > \delta\sqrt{n}\right\}.
\end{aligned}$$

SM5. Auxiliary Lemmas and Proofs.

The following lemma is a generalization of Lemma 11 of [SM1].

LEMMA SM5.1. *Consider any $m \in \mathbb{N}\setminus\{1\}$. For any $V \in \mathcal{V}_m$ and any $z \in \mathbb{C}^n_1$, we have*

$$\frac{1}{n^2}\left\|V^{\mathrm{H}}V - zz^{\mathrm{H}}\right\|_{\mathrm{F}}^2 \leq 2\ell_m(V, z).$$

Proof. Lemma 11 of [SM1] only considers the case where $m = n$. However, its proof holds for any $m \geq 2$, which we include here for completeness. By definition, we have

$$\ell_m(V,z) = 2 - \max_{a\in\mathbb{C}^n:\|a\|^2=1}\left(a^{\mathrm{H}}\left(\frac{1}{n}\sum_{j=1}^n z_jV_j\right) + \left(\frac{1}{n}\sum_{j=1}^n z_jV_j\right)^{\mathrm{H}} a\right) = 2\left(1 - \left\|\frac{1}{n}\sum_{j=1}^n z_jV_j\right\|\right).$$

In addition, we have

$$
\begin{aligned}
n^{-2}\left\|V^{\mathrm{H}}V - zz^{\mathrm{H}}\right\|_{\mathrm{F}}^2 &= \frac{1}{n^2}\sum_{j=1}^n\sum_{l=1}^n |V_j^{\mathrm{H}}V_l - z_j\bar{z}_l|^2 \\
&\leq \frac{1}{n^2}\sum_{j=1}^n\sum_{l=1}^n \left(2 - V_j^{\mathrm{H}}V_l\bar{z}_jz_l - V_l^{\mathrm{H}}V_jz_j\bar{z}_l\right) \\
&= 2\left(1 - \left\|\frac{1}{n}\sum_{j=1}^n z_jV_j\right\|^2\right).
\end{aligned}
$$

Therefore, $n^{-2}\left\|V^{\mathrm{H}}V - zz^{\mathrm{H}}\right\|_{\mathrm{F}}^2 \leq \ell_m(V,z)\left(2 - \frac{1}{2}\ell_m(V,z)\right) \leq 2\ell_m(V,z)$, and the proof is complete. ◻

Proof of Lemma 2.1. We follow the proof of Lemma 12 of [SM2]. We first decompose V and z into orthogonal components:

$$V = a(z^*)^{\mathrm{H}} + \sqrt{n}A \text{ and } z = bz^* + \sqrt{n}\beta, \tag{SM5.1}$$

where $a \in \mathbb{C}^m, A \in \mathbb{C}^{m\times n}, b \in \mathbb{C}, \beta \in \mathbb{C}^n$ and $Az^* = 0, \beta^{\mathrm{H}}z^* = 0$. Note the decomposition on V is always possible as $V = Vz^*(z^*)^{\mathrm{H}} + V(I_n - z^*(z^*)^{\mathrm{H}})$ and $a = Vz^*, \sqrt{n}A = V(I_n - z^*(z^*)^{\mathrm{H}})$. By the definition of the loss ℓ_m in (2.2), there exists some $d \in \mathbb{C}^m$ such $\|d\| = 1$ and $\ell_m(V,z) = n^{-1}\left\|V - dz^{\mathrm{H}}\right\|_{\mathrm{F}}^2$. With the decomposition (SM5.1), it means

$$
\begin{aligned}
n\ell_m(V,z) &= \left\|V - dz^{\mathrm{H}}\right\|_{\mathrm{F}}^2 \\
&= \left\|\left(a(z^*)^{\mathrm{H}} + \sqrt{n}A\right) - d\left(bz^* + \sqrt{n}\beta\right)^{\mathrm{H}}\right\|_{\mathrm{F}}^2 \\
&= \left\|\left(a - d\bar{b}\right)(z^*)^{\mathrm{H}} + \sqrt{n}(A - d\beta^{\mathrm{H}})\right\|_{\mathrm{F}}^2 \\
&= \left\|\left(a - d\bar{b}\right)(z^*)^{\mathrm{H}}\right\|_{\mathrm{F}}^2 + \left\|\sqrt{n}(A - d\beta^{\mathrm{H}})\right\|_{\mathrm{F}}^2 \\
&= n\left\|a - d\bar{b}\right\|^2 + n\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}^2. \qquad \text{(SM5.2)}
\end{aligned}
$$

where the third equation is due to the orthogonality $(A - d\beta^H)z^* = 0$. Then

$$\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}} \leq \sqrt{\ell_m(V,z)}. \tag{SM5.3}$$

We also have

$$
\begin{aligned}
\left\|VY^{\mathrm{H}} - d(Yz)^{\mathrm{H}}\right\|_{\mathrm{F}} &= \left\|V(z^*(z^*)^{\mathrm{H}} + \sigma W)^{\mathrm{H}} - dz^{\mathrm{H}}(z^*(z^*)^{\mathrm{H}} + \sigma W)^{\mathrm{H}}\right\|_{\mathrm{F}} \\
&\leq \left\|(V - dz^{\mathrm{H}})z^*(z^*)^{\mathrm{H}}\right\|_{\mathrm{F}} + \left\|\sigma(V - dz^{\mathrm{H}})W\right\|_{\mathrm{F}} \\
&\leq \left\|(a(z^*)^{\mathrm{H}} - d\bar{b}(z^*)^{\mathrm{H}})z^*(z^*)^{\mathrm{H}}\right\|_{\mathrm{F}} + \sigma\left\|W\right\|\left\|V - dz^{\mathrm{H}}\right\|_{\mathrm{F}} \\
&\leq n\sqrt{n}\left\|a - d\bar{b}\right\| + \sigma\left\|W\right\|\sqrt{n}\sqrt{\ell_m(V,z)}, \qquad \text{(SM5.4)}
\end{aligned}
$$

where the second inequality is due to the fact that $\|B_1B_2\|_{\mathrm{F}} \leq \|B_1\|_{\mathrm{F}}\|B_2\|_{\mathrm{op}}$ for any two matrices B_1, B_2. If

$$\left\|a - d\bar{b}\right\| \leq 6\epsilon\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}} \tag{SM5.5}$$

holds, (SM5.4) and (SM5.3) leads to

$$
\begin{aligned}
\ell_m(VY^{\mathrm{H}}, Yz) &\le \frac{1}{n} \left\| VY^{\mathrm{H}} - d(Yz)^{\mathrm{H}} \right\|_{\mathrm{F}}^2 \\
&\le \frac{1}{n} \left(6\epsilon n\sqrt{n} \left\| A - d\beta^{\mathrm{H}} \right\|_{\mathrm{F}} + \sigma \left\| W \right\| \sqrt{n}\sqrt{\ell_m(V,z)} \right)^2 \\
&\le \frac{1}{n} \left(6\epsilon n\sqrt{n}\sqrt{\ell_m(V,z)} + \sigma \left\| W \right\| \sqrt{n}\sqrt{\ell_m(V,z)} \right)^2 \\
&= n^2 \left(6\epsilon + \frac{\sigma \left\| W \right\|}{n} \right)^2 \ell_m(V,z),
\end{aligned}
$$

which yields the desired result. The remaining proof is devoted to establishing (SM5.5).

Note that

$$
\begin{aligned}
\ell_m(V, z^*) &= \min_{u \in \mathbb{C}^m : \|u\| = 1} n^{-1} \left\| a(z^*)^{\mathrm{H}} + \sqrt{n} A - u(z^*)^{\mathrm{H}} \right\|_{\mathrm{F}}^2 \\
&= \min_{u \in \mathbb{C}^m : \|u\| = 1} n^{-1} \left(\left\| (a-u)(z^*)^{\mathrm{H}} \right\|_{\mathrm{F}}^2 + \left\| \sqrt{n} A \right\|_{\mathrm{F}}^2 \right) \\
&= \min_{u \in \mathbb{C}^m : \|u\| = 1} \left\| a - u \right\|^2 + \left\| A \right\|_{\mathrm{F}}^2 .
\end{aligned}
$$

Since $\ell_m(V, z^*) \le \epsilon^2 < 1/4$, we have $\|A\|_{\mathrm{F}}^2 \le \epsilon^2$, $\|a\| \neq 0$ and $\min_{u \in \mathbb{C}^m : \|u\|=1} \|a-u\|^2 = \|a - a/\|a\|\|^2 = (1 - \|a\|)^2$. Together with $1 = n^{-1} \|V\|_{\mathrm{F}}^2 = n^{-1} \|a(z^*)^{\mathrm{H}}\|^2 + n^{-1} \|\sqrt{n} A\|_{\mathrm{F}}^2 = \|a\|^2 + \|A\|_{\mathrm{F}}^2$, we have

$$
\ell_m(V, z^*) = (1 - \|a\|)^2 + 1 - \|a\|^2 = 2 - 2\|a\| .
$$

Then $\ell_m(V, z^*) \le \epsilon^2$ leads to $1 \ge \|a\| \ge 1 - \epsilon^2/2$. Similarly for z, we have $\|\beta\|^2 \le \epsilon^2$, $1 \ge |b| \ge 1 - \epsilon^2/2$ and $1 = |b|^2 + \|\beta\|^2$. Since $\epsilon < 1/2$, we have $\|a\| + |b| > 1$, and consequently $|\|a\| - |b|| \le |\|a\| - |b|| (\|a\| + |b|) = |\|a\|^2 - |b|^2|$. Since $\|a\|^2 + \|A\|_{\mathrm{F}}^2 = |b|^2 + \|\beta\|^2$, we have $|\, \|a\|^2 - |b|^2| = |\, \|\beta\|^2 - \|A\|_{\mathrm{F}}^2 \,|$. Together with $\|A\|_{\mathrm{F}}^2, \|\beta\|^2 \le \epsilon^2$, we have

$$
\begin{aligned}
|\|a\| - |b|| &\le |\, \|\beta\|^2 - \|A\|_{\mathrm{F}}^2 \,| = |\, \|\beta\| - \|A\|_{\mathrm{F}} \,| \left(\|\beta\| + \|A\|_{\mathrm{F}} \right) \\
&\le 2\epsilon |\, \|\beta\| - \|A\|_{\mathrm{F}} \,| \le 2\epsilon \left\| A - d\beta^{\mathrm{H}} \right\|_{\mathrm{F}} . \qquad \text{(SM5.6)}
\end{aligned}
$$

Note that

$$
\begin{aligned}
\left\| a - d\bar{b} \right\| &= \left\| a - \frac{a}{\|a\|}|b| + \frac{a}{\|a\|}\frac{b}{|b|}\bar{b} - d\bar{b} \right\| \\
&\le \left\| a - \frac{a}{\|a\|}|b| \right\| + \left\| \left(\frac{a}{\|a\|}\frac{b}{|b|} - d \right) \bar{b} \right\| \\
&= |\|a\| - |b|| + \left\| \frac{a}{\|a\|}\frac{b}{|b|} - d \right\| |b| \\
&\le 2\epsilon \left\| A - d\beta^{\mathrm{H}} \right\|_{\mathrm{F}} + \left\| \frac{a}{\|a\|}\frac{b}{|b|} - d \right\| ,
\end{aligned}
$$

where in the last inequality we use $|b| \le 1$. Hence, to establish (SM5.5), we only need to show

$$
\left\| \frac{a}{\|a\|}\frac{b}{|b|} - d \right\| \le 4\epsilon \left\| A - d\beta^{\mathrm{H}} \right\|_{\mathrm{F}} . \qquad \text{(SM5.7)}
$$

To prove (SM5.7), define $d_0 = \frac{a}{\|a\|}\frac{b}{|b|} \in \mathbb{C}^m$. Then $\|d_0\| = 1$. Similar to (SM5.2), we have $\|V - d_0 z^{\mathrm{H}}\|_{\mathrm{F}}^2 = n\left\|a - d_0\bar{b}\right\|^2 + n\left\|A - d_0\beta^{\mathrm{H}}\right\|_{\mathrm{F}}^2$. By the definition of d, $\|V - dz^{\mathrm{H}}\|_{\mathrm{F}}^2 \le \|V - d_0 z^{\mathrm{H}}\|_{\mathrm{F}}^2$, which leads to

$$\left\|a - d\bar{b}\right\|^2 + \left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}^2 \le \left\|a - d_0\bar{b}\right\|^2 + \left\|A - d_0\beta^{\mathrm{H}}\right\|_{\mathrm{F}}^2.$$

Note that $d_0\bar{b} = a\frac{|b|}{\|a\|}$ is proportional to a and $\left\|d_0\bar{b}\right\| = \left\|d\bar{b}\right\| = |b|$. Let $\theta \in [0,\pi]$ be the angle between a and $d\bar{b}$ in $\mathbb{C}^m$. By the cosine formula of triangles, we have

$$\left\|a - d\bar{b}\right\|^2 = \|a\|^2 + \left\|d\bar{b}\right\|^2 - 2\|a\|\left|d\bar{b}\right|\cos(\theta) = \|a\|^2 + |b|^2 - 2\|a\|\,|b|\cos(\theta)$$

$$\left\|a - d_0\bar{b}\right\|^2 = \left\|a - a\frac{|b|}{\|a\|}\right\|^2 = \|a\|^2 + |b|^2 - 2\|a\|\,|b|$$

$$\text{and } \|d - d_0\|^2 = \|d\|^2 + \|d_0\|^2 - 2\|d\|\,\|d_0\|\cos(\theta) = 2(1 - \cos(\theta)). \tag{SM5.8}$$

Hence, $\left\|a - d\bar{b}\right\|^2 - \left\|a - d_0\bar{b}\right\|^2 = 2\|a\|\,|b|(1 - \cos(\theta))$. By the triangle inequality, $\left\|A - d_0\beta^{\mathrm{H}}\right\|_{\mathrm{F}} - \left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}} \le \left\|(d_0 - d)\beta^{\mathrm{H}}\right\|_{\mathrm{F}} = \|d_0 - d\|\,\|\beta\| \le \epsilon\|d_0 - d\|$ where in the last inequality we use $\|\beta\| \le \epsilon$. Then,

$$\begin{aligned}
2\|a\|\,|b|(1 - \cos(\theta)) &\le \left\|A - d_0\beta^{\mathrm{H}}\right\|_{\mathrm{F}}^2 - \left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}^2 \\
&= \left(\left\|A - d_0\beta^{\mathrm{H}}\right\|_{\mathrm{F}} - \left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}\right)\left(\left\|A - d_0\beta^{\mathrm{H}}\right\|_{\mathrm{F}} - \left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}} + 2\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}\right) \\
&\le \epsilon\|d_0 - d\|\left(\epsilon\|d_0 - d\| + 2\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}\right).
\end{aligned}$$

By (SM5.8), it becomes $\|a\|\,|b|\,\|d_0 - d\|^2 \le \epsilon\|d_0 - d\|\left(\epsilon\|d_0 - d\| + 2\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}\right)$, which further leads to

$$\left(\epsilon^{-1}\|a\|\,|b| - \epsilon\right)\|d_0 - d\| \le 2\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}.$$

Since $\|a\|, |b| \ge 1 - \epsilon^2/2$, we have $\epsilon^{-1}\|a\|\,|b| - \epsilon \ge \epsilon^{-1}(1 - \epsilon^2/2)^2 - \epsilon \ge \epsilon^{-1}(1 - \epsilon^2) - \epsilon = \epsilon^{-1}(1 - 2\epsilon^2) > (2\epsilon)^{-1}$ where the last inequality is due to $\epsilon < 1/2$. Hence, $(2\epsilon)^{-1}\|d_0 - d\| \le 2\left\|A - d\beta^{\mathrm{H}}\right\|_{\mathrm{F}}$, which establishes (SM5.7). The proof of the lemma is complete. ∎

REFERENCES

[SM1] C. Gao and A. Y. Zhang, *Sdp achieves exact minimax optimality in phase synchronization*, IEEE Transactions on Information Theory, 68 (2022), pp. 5374–5390.

[SM2] Y. Zhong and N. Boumal, *Near-optimal bounds for phase synchronization*, SIAM Journal on Optimization, 28 (2018), pp. 989–1016.