# Supplementary Material

Baoshun Shi[a,b], Zheng Liu[a,b], Kexun Liu[a,b], Yueming Su[c]

[a] *School of Information Science and Engineering, Yanshan University, Qinhuangdao, 066004, Hebei Province, China*
[b] *Hebei Key Laboratory of Information Transmission and Signal Processing, Qinhuangdao, 066004, Hebei Province, China*
[c] *School of Information, Beijing Wuzi University, Beijing, 101149, China*

## 1. Theorem 1:

**Theorem 1.** *For any $\boldsymbol{x} \in \mathbb{R}^N$ whose elements admit $\boldsymbol{x}_i \in [0, 1]$ and some universal constant L independent of M, PDSNet is bounded such that*

$$\frac{1}{M} \left\| \boldsymbol{x} - \boldsymbol{W}_1^{\mathrm{T}} \boldsymbol{W}_2^{\mathrm{T}} T \left[ \boldsymbol{W}_2 T \left( \boldsymbol{W}_1 \boldsymbol{x}, \boldsymbol{e}_1 \right), \boldsymbol{e}_2 \right] \right\|_2^2 \leq \sigma^2 L. \tag{A.1}$$

**Proof**: In Eqn. (A.1), $\boldsymbol{e}_1 \in \mathbb{R}^M$ and $\boldsymbol{e}_2 \in \mathbb{R}^M$ represent the threshold vectors whose elements are utilized for shrinking $\boldsymbol{W}_1\boldsymbol{x}$ and $\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1)$ respectively. $\boldsymbol{W}_1 \in \mathbb{R}^{M\times N}$ and $\boldsymbol{W}_2 \in \mathbb{R}^{M\times N}$ are tight frames satisfying the tight property $\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{W}_1 = \boldsymbol{I}$ and $\boldsymbol{W}_2^{\mathrm{T}}\boldsymbol{W}_2 = \boldsymbol{I}$ respectively. Therefore, we have

$$\begin{aligned}
&\left\| \boldsymbol{x} - \boldsymbol{W}_1^{\mathrm{T}} \boldsymbol{W}_2^{\mathrm{T}} T \left[ \boldsymbol{W}_2 T \left( \boldsymbol{W}_1 \boldsymbol{x}, \boldsymbol{e}_1 \right), \boldsymbol{e}_2 \right] \right\|_2^2 \\
&= \left\| \boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_1^{\mathrm{T}} \boldsymbol{W}_2^{\mathrm{T}} T \left[ \boldsymbol{W}_2 T \left( \boldsymbol{W}_1 \boldsymbol{x}, \boldsymbol{e}_1 \right), \boldsymbol{e}_2 \right] \right\|_2^2 \\
&\leq \left\| \boldsymbol{W}_1^{\mathrm{T}} \right\|_2^2 \left\| \boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2^{\mathrm{T}} T \left[ \boldsymbol{W}_2 T \left( \boldsymbol{W}_1 \boldsymbol{x}, \boldsymbol{e}_1 \right), \boldsymbol{e}_2 \right] \right\|_2^2 \\
&= \left\| \boldsymbol{W}_1^{\mathrm{T}} \right\|_2^2 \left\| \boldsymbol{W}_2^{\mathrm{T}}\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2^{\mathrm{T}} T \left[ \boldsymbol{W}_2 T \left( \boldsymbol{W}_1 \boldsymbol{x}, \boldsymbol{e}_1 \right), \boldsymbol{e}_2 \right] \right\|_2^2 \\
&\leq \left\| \boldsymbol{W}_1^{\mathrm{T}} \right\|_2^2 \left\| \boldsymbol{W}_2^{\mathrm{T}} \right\|_2^2 \left\| \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T \left[ \boldsymbol{W}_2 T \left( \boldsymbol{W}_1 \boldsymbol{x}, \boldsymbol{e}_1 \right), \boldsymbol{e}_2 \right] \right\|_2^2 .
\end{aligned} \tag{A.2}$$

Due to $\left\| \boldsymbol{W}_1^{\mathrm{T}} \right\|_2^2 = \lambda_{max}(\boldsymbol{W}_1\boldsymbol{W}_1^{\mathrm{T}})$ where $\lambda_{max}(\bullet)$ represents the maximum eigenvalue. We assume that the maximum eigenvalue of the $\boldsymbol{W}_1\boldsymbol{W}_1^{\mathrm{T}}$ is $\lambda_{1max}$. Following similar steps, $\left\| \boldsymbol{W}_2^{\mathrm{T}} \right\|_2^2 = \lambda_{max}(\boldsymbol{W}_2\boldsymbol{W}_2^{\mathrm{T}})$ , and we assume that the maximum

*Corresponding author
*Email address:* `shibaoshun@ysu.edu.cn` (Baoshun Shi)

eigenvalue of the $\boldsymbol{W}_2\boldsymbol{W}_2^{\mathrm{T}}$ is $\lambda_{2max}$. Therefore, the upper bound of Eqn. (A.2) can be further determined as follows:

$$\begin{aligned} &\left\|\boldsymbol{x} - \boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{W}_2^{\mathrm{T}}T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\right\|_2^2 \\ \leq\ &\lambda_{1max}\lambda_{2max}\left\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\right\|_2^2 . \end{aligned} \tag{A.3}$$

Let $(\boldsymbol{W}_l\boldsymbol{x})_i, l = 1, 2$ represent the $i$th element of $\boldsymbol{W}_l\boldsymbol{x}$ and $\varepsilon_{li}$ which is non-negative denote the $i$th element of $\boldsymbol{e}_l$. The soft thresholding operator $T\left[(\boldsymbol{W}_l\boldsymbol{x})_i, \varepsilon_{li}\right], l = 1, 2$ is defined as

$$T\left[(\boldsymbol{W}_l\boldsymbol{x})_i, \varepsilon_{li}\right] = \begin{cases} (\boldsymbol{W}_l\boldsymbol{x})_i + \varepsilon_{li}, & (\boldsymbol{W}_l\boldsymbol{x})_i < -\varepsilon_{li} \\ 0, & |(\boldsymbol{W}_l\boldsymbol{x})_i| \leqslant \varepsilon_{li} \\ (\boldsymbol{W}_l\boldsymbol{x})_i - \varepsilon_{li}, & (\boldsymbol{W}_l\boldsymbol{x})_i > \varepsilon_{li} \end{cases} . \tag{A.4}$$

We consider the first shrinking process $T\left[\boldsymbol{W}_1\boldsymbol{x}, e_1\right]$, one of the following situations will happen.

**S1:** Any element of $\boldsymbol{W}_1\boldsymbol{x}$ satisfies $(\boldsymbol{W}_1\boldsymbol{x})_i < -\varepsilon_{1i}$, then $T\left[\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right] = \boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1$. Therefore,

$$T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] = T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]. \tag{A.5}$$

**S2:** Any element of $\boldsymbol{W}_1\boldsymbol{x}$ satisfies $|(\boldsymbol{W}_1\boldsymbol{x})_i| \leq \varepsilon_{1i}$, then $T\left[\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right] = 0$. Therefore,

$$T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] = T\left[0, \boldsymbol{e}_2\right] = 0. \tag{A.6}$$

**S3:** Any element of $\boldsymbol{W}_1\boldsymbol{x}$ satisfies $(\boldsymbol{W}_1\boldsymbol{x})_i > \varepsilon_{1i}$, then $T\left[\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right] = \boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1$. Therefore,

$$T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] = T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]. \tag{A.7}$$

**S4:** Any two or all three of the above situations occur.

When **S1** happens, according to the **Lemma 2**, we have

$$\begin{aligned} & \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\ \leq & \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 + 2\|\boldsymbol{e}_1\|_2\|\boldsymbol{e}_2\|_2. \end{aligned} \tag{A.8}$$

When **S2** happens, according to the **Lemma 3**, we have

$$\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \leq \|\boldsymbol{e}_1\|_2^2. \tag{A.9}$$

When **S3** happens, according to the **Lemma 4**, we have

$$\begin{aligned} & \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\ \leq & \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 + 2\|\boldsymbol{e}_1\|_2\|\boldsymbol{e}_2\|_2. \end{aligned} \tag{A.10}$$

**S4** is a union of the **S1**, **S2** and **S3**. Therefore, as long as we find the upper bound of $\|\boldsymbol{W}_2\boldsymbol{W}_1 x - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2$ under the **S1**-and **S3**, the upper bound of that under the **S4** will be determined. Hence, when **S4** happens, the upper bound of $\|\boldsymbol{W}_2\boldsymbol{W}_1 x - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2$ is the maximum upper bound of the first three situations. Based on Eqn. (A.8), Eqn. (A.9), and Eqn. (A.10), we have

$$\begin{aligned} & \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\ \leq & \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 + 2\|\boldsymbol{e}_1\|_2\|\boldsymbol{e}_2\|_2. \end{aligned} \tag{A.11}$$

Recall the definition of threshold vectors $\boldsymbol{e}_l = \boldsymbol{c}_l \odot \boldsymbol{m}$, $l = 1, 2$, and each element of proportional constant vectors $\boldsymbol{c}_l$ has a limited range $[\boldsymbol{c}_{\min}, \boldsymbol{c}_{max}]$. Let $\varepsilon_{1max} = c_{1max} \cdot \sigma$ and $\varepsilon_{2max} = c_{2max} \cdot \sigma$ denote the maximum element of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ respectively. Thus,

$$\begin{aligned} & \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 + 2\,\|\boldsymbol{e}_1\|_2\,\|\boldsymbol{e}_2\|_2 \\ \leq & M\varepsilon_{1max}^2 + M\varepsilon_{2max}^2 + 2M\varepsilon_{1max}\varepsilon_{2max} \\ \leq & Mc_{1max}^2\sigma^2 + Mc_{2max}^2\sigma^2 + 2Mc_{1max}c_{2max}\sigma^2 \end{aligned} \tag{A.12}$$

According to Eqn. (A.3), Eqn. (A.11), and Eqn. (A.12), we can get

$$\begin{aligned}
&\left\|\boldsymbol{x} - \boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{W}_2^{\mathrm{T}} T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\right\|_2^2 \\
\leq& \lambda_{1max}\lambda_{2max} M\sigma^2 \left(c_{1max}^2 + c_{2max}^2 + 2c_{1max}c_{2max}\right) \\
\leq& M\sigma^2 L.
\end{aligned} \tag{A.13}$$

Here, $L = \lambda_{1max}\lambda_{2max}\left(c_{1max}^2 + c_{2max}^2 + 2c_{1max}c_{2max}\right)$.

This completes the proof.

## 2. Lemma 2

**Lemma 2.** *When **S1** happens, we have*

$$\begin{aligned}
&\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\
\leq& \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 + 2\|\boldsymbol{e}_1\|_2\|\boldsymbol{e}_2\|_2 .
\end{aligned} \tag{1}$$

**Proof:** Let $[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1)]_i$ represent the $i$th element of $\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1)$ and $\varepsilon_{2i}$ denote the $i$th element of $\boldsymbol{e}_2$ . We consider the second shrinking process $T\left[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1), \boldsymbol{e}_2\right]$ and the following cases must happen.

Case 1: Any element of $\boldsymbol{W}_2(\boldsymbol{W}_1 x + \boldsymbol{e}_1)$ satisfies $[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1)]_i < -\varepsilon_{2i}$, then

$$\begin{aligned}
&T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\
=& T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\
=& \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2.
\end{aligned} \tag{2}$$

Based on Eqn. (2), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2$ can be rewritten as

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{e}_1 - \boldsymbol{e}_2||_2^2 \\
&= ||-\boldsymbol{W}_2\boldsymbol{e}_1 - \boldsymbol{e}_2||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2\,||\boldsymbol{e}_1||_2\,||\boldsymbol{e}_2||_2\,.
\end{aligned} \tag{3}$$

Case 2: Any element of $\boldsymbol{W}_2\left(\boldsymbol{W}_1 x + \boldsymbol{e}_1\right)$ satisfies $|[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1\right)]_i| \leqslant \varepsilon_{2i}$, then

$$T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] = T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] = 0 \tag{4}$$

Based on Eqn. (4), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2$ can be rewritten as

$$||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2 = ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2\,. \tag{5}$$

Since $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1||_2^2 \geq ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2 - ||\boldsymbol{W}_2\boldsymbol{e}_1||_2^2$ and $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1||_2^2 \leq ||\boldsymbol{e}_2||_2^2$, it follows that

$$||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2 \leq ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1||_2^2 + ||\boldsymbol{W}_2\boldsymbol{e}_1||_2^2 \leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2\,. \tag{6}$$

Case 3: Any element of $\boldsymbol{W}_2\left(\boldsymbol{W}_1 x + \boldsymbol{e}_1\right)$ satisfies $[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1\right)]_i > \varepsilon_{2i}$, then

$$\begin{aligned}
&T\left[\boldsymbol{W}_2T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\
=&T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\
=&\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1 - \boldsymbol{e}_2.
\end{aligned} \tag{7}$$

Based on Eqn. (7), $\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2$ can be rewritten as

$$\begin{aligned}
&\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\
&= \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2\|_2^2 \\
&= \|-\boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2\|_2^2 \\
&\leq \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 .
\end{aligned} \tag{8}$$

In addition to the Case1-Case3, another case is the union of the above three cases. In this case, the upper bound of $\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2$ is the maximum upper bound of the above three cases. Hence, Eqn. (1) holds under any case, i.e.,

$$\begin{aligned}
&\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\
&= \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\
&\leq \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 + 2\|\boldsymbol{e}_1\|_2\|\boldsymbol{e}_2\|_2 .
\end{aligned} \tag{9}$$

This completes the proof.

## 3. Lemma 3

**Lemma 3.** *When **S2** happens, we have*

$$\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \leq \|\boldsymbol{e}_1\|_2^2 . \tag{10}$$

**Proof:** Since under the **S2**, we have $|(\boldsymbol{W}_1\boldsymbol{x})_i| \leq \varepsilon_{1i}$, it follows that $\|\boldsymbol{W}_1\boldsymbol{x}\|_2^2 \leq \|\boldsymbol{e}_1\|_2^2$. Consequently, we have

$$\begin{aligned}
&\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\
&= \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}\|_2^2 \\
&\leq \|\boldsymbol{e}_1\|_2^2 .
\end{aligned} \tag{11}$$

This completes the proof.

## 4. Lemma 4

**Lemma 4.** *When **S3** happens, we have*

$$\begin{aligned} & \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\ \leq & \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 + 2\|\boldsymbol{e}_1\|_2\|\boldsymbol{e}_2\|_2 . \end{aligned} \tag{12}$$

**Proof:** Let $[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1)]_i$ represent the $i$th element of $\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1)$ and $\varepsilon_{2i}$ denote the $i$th element of $\boldsymbol{e}_2$. We consider the second shrinking process $T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1), \boldsymbol{e}_2]$ and the following cases must happen.

Case 1: Any element of $\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1)$ satisfies $[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1)]_i < -\varepsilon_{2i}$, then

$$\begin{aligned} & T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\ = & T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\ = & \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2. \end{aligned} \tag{13}$$

Based on Eqn. (13), $\|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]\|_2^2$ can be rewritten as

$$\begin{aligned} & \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]\|_2^2 \\ = & \|\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1 - \boldsymbol{e}_2\|_2^2 \\ = & \|\boldsymbol{W}_2\boldsymbol{e}_1 - \boldsymbol{e}_2\|_2^2 \\ \leq & \|\boldsymbol{e}_1\|_2^2 + \|\boldsymbol{e}_2\|_2^2 \end{aligned} \tag{14}$$

Case 2: Any element of $\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1)$ satisfies $|[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1)]_i| \leq \varepsilon_{2i}$, then

$$T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] = T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] = 0 \tag{15}$$

Based on Eqn. (15), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2$ can be rewritten as

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2
\end{aligned} \tag{16}$$

Case 3: Any element of $\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1\right)$ satisfies $\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1\right)\right]_i > \varepsilon_{2i}$, then

$$\begin{aligned}
&T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\
=&T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1\right), \boldsymbol{e}_2\right] \\
=&\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{e}_1 - \boldsymbol{e}_2
\end{aligned} \tag{17}$$

Based on Eqn. (17), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2$ can be rewritten as

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2\,||\boldsymbol{e}_1||_2\,||\boldsymbol{e}_2||_2
\end{aligned} \tag{18}$$

Similar to **Lemma 2**, besides the Case1-Case3, another case is the union of the above three cases. In this case, the upper bound of $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2$ is the maximum upper bound of the above three cases. Hence, Eqn. (25) holds under any case, i.e.,

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2 T\left(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T\left[\boldsymbol{W}_2\left(\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{e}_1\right), \boldsymbol{e}_2\right]||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2\,||\boldsymbol{e}_1||_2\,||\boldsymbol{e}_2||_2
\end{aligned} \tag{19}$$

This completes the proof.