1. SUPPLEMENTARY MATERIAL

1.1. Proof of Theorem 1

Theorem 1: *For any $\boldsymbol{x} \in \mathbb{R}^N$ whose element admits $x_i \in [0,1]$, and each element of proportional constant $\boldsymbol{c_l}$ has a limited range $c_i \in [c_{min}, c_{max}]$. For some universal constant L independent of M and noise level σ, DoubleTFCnet is a bounded denoiser such that*

$$\frac{1}{M}\|\boldsymbol{x} - \boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{W}_2^{\mathrm{T}}T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]\|_2^2 \leq \sigma^2 L. \quad (1)$$

Proof: In Eqn. (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 $\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1)$ respectively. $\boldsymbol{W}_1 \in \mathbb{R}^{M\times N}$ and $\boldsymbol{W}_2 \in \mathbb{R}^{M\times M}$ 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

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

Due to $||\boldsymbol{W}_1^{\mathrm{T}}||_2^2 = \lambda_{max}(\boldsymbol{W}_1{\boldsymbol{W}_1}^{\mathrm{T}})$ 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 λ_{1max}. Following similar steps, $||\boldsymbol{W}_2^{\mathrm{T}}||_2^2 = \lambda_{max}(\boldsymbol{W}_2{\boldsymbol{W}_2}^{\mathrm{T}})$, and we assume that the maximum eigenvalue of the $\boldsymbol{W}_2{\boldsymbol{W}_2}^{\mathrm{T}}$ is λ_{2max}. Therefore, the upper bound of Eqn. (2) can be further determined as follows:

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

Let $(\boldsymbol{W}_1\boldsymbol{x})_i$ represent the *i*-th element of $\boldsymbol{W}_1\boldsymbol{x}$ and ε_{1i} denote the *i*-th element of $\boldsymbol{e}_1$. The soft thresholding operator $T[(\boldsymbol{W}_l\boldsymbol{x})_i, \varepsilon_{li}], l = 1, 2$ is defined as

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

We consider the first shrinking process $T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1)$, 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 $T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1) = \boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{e}_1$. Therefore,

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

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

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

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

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

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

When **S1** happens, according to the *Lemma 1*, we have

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2.
\end{aligned} \quad (8)$$

When **S2** happens, according to the *Lemma 2*, we have

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2.
\end{aligned} \quad (9)$$

When **S3** happens, according to the *Lemma 3*, we have

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2.
\end{aligned} \quad (10)$$

S4 is a union of the **S1**, **S2** and **S3**. Therefore, as long as we find the upper bound of $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2$ under the **S1**-**S3**, the upper bound of that under the **S4** will be determined. Hence, when **S4** happens, the upper bound of $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2$ is the maximum upper bound of the first three situations. Based on Eqn. (8), Eqn. (9) and Eqn. (10), we have

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2.
\end{aligned} \quad (11)$$

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 $c_{li} \in [c_{min}, c_{max}]$. 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,

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

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

$$\begin{aligned}
&||\boldsymbol{x} - \boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{W}_2^{\mathrm{T}}T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&\leq \lambda_{1max}\lambda_{2max}M\sigma^2(c_{1max}^2 + c_{2max}^2 + 2c_{1max}c_{2max}) \\
&\leq M\sigma^2 L.
\end{aligned} \quad (13)$$

Here, $L = \lambda_{1max}\lambda_{2max}(c_{1max}^2 + c_{2max}^2 + 2c_{1max}c_{2max})$. This completes the proof.

Lemma 1: When **S1** happens, we have

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2.
\end{aligned} \quad (14)$$

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 ε_{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.

Case1: 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

$$\begin{aligned} & T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{W}_2\boldsymbol{e}_1+\boldsymbol{e}_2. \end{aligned} \tag{15}$$

Based on Eqn. (15), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2$ can be rewritten as

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{W}_2\boldsymbol{e}_1-\boldsymbol{e}_2||_2^2 \\ &= ||-\boldsymbol{W}_2\boldsymbol{e}_1-\boldsymbol{e}_2||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2+||\boldsymbol{e}_2||_2^2+2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2. \end{aligned} \tag{16}$$

Case2: Any element of $\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{e}_1)$ satisfies $|[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{e}_1)]_i| \le \varepsilon_{2i}$, then

$$\begin{aligned} & T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= \boldsymbol{0}. \end{aligned} \tag{17}$$

Based on Eqn. (17), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2$ can be rewritten as

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2. \end{aligned} \tag{18}$$

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

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2 \\ &\le ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{W}_2\boldsymbol{e}_1||_2^2+||\boldsymbol{W}_2\boldsymbol{e}_1||_2^2 \\ &\le ||\boldsymbol{e}_2||_2^2+||\boldsymbol{e}_1||_2^2. \end{aligned} \tag{19}$$

Case3: 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

$$\begin{aligned} & T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{W}_2\boldsymbol{e}_1-\boldsymbol{e}_2. \end{aligned} \tag{20}$$

Based on Eqn. (20), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2$ can be rewritten as

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{W}_2\boldsymbol{e}_1+\boldsymbol{e}_2||_2^2 \\ &= ||-\boldsymbol{W}_2\boldsymbol{e}_1+\boldsymbol{e}_2||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2+||\boldsymbol{e}_2||_2^2. \end{aligned} \tag{21}$$

In addition to the *Case1*-*Case3*, another case is the union of the above three cases. In this case, the upper bound of $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2$ is the maximum upper bound of the above three cases. Hence, Eqn. (14) holds under any case, i.e.,

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2+||\boldsymbol{e}_2||_2^2+2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2. \end{aligned} \tag{22}$$

Lemma 2: When **S2** happens, we have

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2. \end{aligned} \tag{23}$$

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

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2. \end{aligned} \tag{24}$$

Lemma 3: When **S3** happens, we have

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2+||\boldsymbol{e}_2||_2^2+2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2. \end{aligned} \tag{25}$$

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 ε_{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.

Case1: 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

$$\begin{aligned} & T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{W}_2\boldsymbol{e}_1+\boldsymbol{e}_2. \end{aligned} \tag{26}$$

Based on Eqn. (26), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2$ can be rewritten as

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{W}_2\boldsymbol{e}_1-\boldsymbol{e}_2||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{e}_1-\boldsymbol{e}_2||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2+||\boldsymbol{e}_2||_2^2. \end{aligned} \tag{27}$$

Case2: Any element of $\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{e}_1)$ satisfies $|[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{e}_1)]_i| \le \varepsilon_{2i}$, then

$$\begin{aligned} & T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{e}_1),\boldsymbol{e}_2] \\ &= \boldsymbol{0}. \end{aligned} \tag{28}$$

Based on Eqn. (28), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2$ can be rewritten as

$$\begin{aligned} & ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}-T[\boldsymbol{W}_2T(\boldsymbol{W}_1\boldsymbol{x},\boldsymbol{e}_1),\boldsymbol{e}_2]||_2^2 \\ &= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x}||_2^2 \\ &\le ||\boldsymbol{e}_1||_2^2+||\boldsymbol{e}_2||_2^2. \end{aligned} \tag{29}$$

Case3: 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

$$\begin{aligned}
&T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2] \\
&= T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{e}_1), \boldsymbol{e}_2] \\
&= \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{e}_1 - \boldsymbol{e}_2.
\end{aligned} \tag{30}$$

Based on Eqn. (30), $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2$ can be rewritten as

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - \boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} + \boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{e}_1 + \boldsymbol{e}_2||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2.
\end{aligned} \tag{31}$$

Similar to Lemma 1, besides the *Case1*-*Case3*, another case is the union of the above three cases. In this case, the upper bound of $||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2$ is the maximum upper bound of the above three cases. Hence, Eqn. (25) holds under any case, i.e.,

$$\begin{aligned}
&||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2 T(\boldsymbol{W}_1\boldsymbol{x}, \boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&= ||\boldsymbol{W}_2\boldsymbol{W}_1\boldsymbol{x} - T[\boldsymbol{W}_2(\boldsymbol{W}_1\boldsymbol{x}-\boldsymbol{e}_1), \boldsymbol{e}_2]||_2^2 \\
&\leq ||\boldsymbol{e}_1||_2^2 + ||\boldsymbol{e}_2||_2^2 + 2||\boldsymbol{e}_1||_2||\boldsymbol{e}_2||_2.
\end{aligned} \tag{32}$$