Lemma 1. *For $x \in \mathbb{R}^2$, the 2D shrinkage function $s^l : \mathbb{R}^2 \to \mathbb{R}^2$ defined as*

$$s^l(\mathbf{x}) = s_{\frac{1}{\beta^l}}(\mathbf{x}) \triangleq \max\left\{\|\mathbf{x}\|_2 - \frac{1}{\beta^l}, 0\right\} \frac{\mathbf{x}}{\|\mathbf{x}\|_2}, \tag{1}$$

where $\beta^l > 0$, is nonexpansive

$$\|s^l(\mathbf{x}_1) - s^l(\mathbf{x}_2)\|_2 \leq \|\mathbf{x}_1 - \mathbf{x}_2\|_2. \tag{2}$$

Furthermore, if $\|s^l(\mathbf{x}_1) - s^l(\mathbf{x}_2)\|_2 = \|\mathbf{x}_1 - \mathbf{x}_2\|_2$, then $s^l(\mathbf{x}_1) - s^l(\mathbf{x}_2) = \mathbf{x}_1 - \mathbf{x}_2$.

Proof. Refer to Proposition 3.1 in [R1]. □

Lemma 2. *Assume $\mathcal{N}(\mathbf{K}) \cap \mathcal{N}(\mathbf{D}^*) = 0$, where $\mathcal{N}(\cdot)$ represents the null space of a matrix. Then for any $\mathbf{w} \neq \widetilde{\mathbf{w}} \in \mathbb{R}^n$, it holds that*

$$\|h^l(\mathbf{w}) - h^l(\widetilde{\mathbf{w}})\|_2 \leq \|\mathbf{w} - \widetilde{\mathbf{w}}\|_2. \tag{3}$$

Proof. Refer to Proposition 3.2 in [R2]. □

Lemma 3. *Let $\{a_k\}_k$ be a real non-negative sequence. If $a_{k+1} \leq a_k + \epsilon_k$ for all k, $\epsilon_k \geq 0$ and $\sum_{i=1}^{\infty} \epsilon_k$ converges, then a_k also converges.*

Proof. It is easy to redsee that a_k is bounded. Therefore, $\exists \mathcal{K} \subset \mathbb{N}$ s.t. $\{a_k\}_{k\in\mathcal{K}}$ converges. redFor all $\varepsilon > 0$, choose $N_1 \in \mathbb{N}$ s.t. $|a_{n_1} - a_{n_2}| < \frac{\varepsilon}{2}, \forall n_1, n_2 \in \mathcal{K}$ and red$n_1, n_2 > N_1$. Now that $\sum_{i=1}^{k} \epsilon_i$ converges, pick $N_2 \in \mathbb{N}$ s.t. $\sum_{i=N_2+1}^{\infty} \epsilon_i < \frac{\varepsilon}{2}$. Let $N = \min\{n \in \mathcal{K} : n > \max\{N_1, N_2\}\}$. redFor all $m \geq k > N$, choose $k_1 = \max\{n \in \mathcal{K} : n < k\}$, $k_2 = \min\{n \in \mathcal{K} : n > m\}$, then $k_1 > \max\{N_1, N_2\}$, $|a_{k_1} - a_{k_2}| < \frac{\varepsilon}{2}$ and

$$a_k - a_{k_1} \leq \sum_{i=k_1}^{k-1} \epsilon_i, \tag{4}$$

$$a_{k_2} - a_m \leq \sum_{i=m}^{k_2-1} \epsilon_i. \tag{5}$$

Adding (4) and (5) yields $a_k - a_m + a_{k_2} - a_{k_1} \leq \sum_{i=k_1}^{k-1} \epsilon_i + \sum_{i=m}^{k_2-1} \epsilon_i \leq \sum_{i=k_1}^{\infty} \epsilon_i \leq \sum_{i=N2+1}^{\infty} \epsilon_i < \frac{\varepsilon}{2}$, hence

$$a_k - a_m \leq |a_{k_1} - a_{k_2}| + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{6}$$

On the other hand, $k > k_1 > N_2$ and

$$a_m - a_k \leq \sum_{i=k}^{m-1} \epsilon_i \leq \sum_{i=N2+1}^{\infty} \epsilon_i < \varepsilon. \tag{7}$$

Combining (6) and (7) redgives $|a_k - a_m| < \varepsilon$, which implies $\{a_k\}_k$ is a red Cauchy sequence, and hence it converges. □

Theorem 1. *(Convergence result) For a fixed $\mu > 0$, suppose $\mathbf{D}^l = \mathbf{D}^* + \xi_l \mathbf{E}^l$ and $\beta^l = \beta^* + \gamma_l$, where $\mathbf{E}^l$ is a Toeplitz matrix with the same structure as $\mathbf{D}^l$, $\mathbf{D}^* \in \mathbb{R}^{m\times n}$ and $\beta^* \in \mathbb{R}$ are fixed quantities, and $\xi_l \in \mathbb{R}$, $\gamma_l \in \mathbb{R}$ are absolutely summable, redmeaning $\sum_l |\xi_l|$ and $\sum_l |\gamma_l|$ both converge. Then the sequence $\{(\mathbf{w}^l, \mathbf{u}^l)\}$ generated from any starting point $\{(\mathbf{w}^0, \mathbf{u}^0)\}$ converges to a solution $\{(\mathbf{w}^*, \mathbf{u}^*)\}$ of (9) (in the original manuscript) and (10) (in the original manuscript) in which both $\mathbf{D}^l$ and β^l converge to $\mathbf{D}^*$ and β^* respectively.*

Proof. According to the assumption on $\mathbf{D}^l$ and β^l, we have $\lim_{l\to+\infty}\mathbf{D}^l = \lim_{l\to+\infty}(\mathbf{D}^*+\xi_l\mathbf{E}^l) = \mathbf{D}^*$ and $\lim_{l\to+\infty}\beta^l = \lim_{l\to+\infty}(\beta^*+\gamma_l) = \beta^*$. It therefore follows that $\lim_{l\to+\infty}\mathbf{M}^l = \lim_{l\to+\infty}(\mathbf{D}^l)^T\mathbf{D}^l + \frac{\mu}{\beta^l}\mathbf{K}^T\mathbf{K} = (\mathbf{D}^*)^T\mathbf{D}^* + \frac{\mu}{\beta^*}\mathbf{K}^T\mathbf{K} = \mathbf{M}^*$. By continuity of s^l and h^l, $\forall \mathbf{x}\in\mathbb{R}^n$, we have $\lim_{l\to+\infty}s^l(\mathbf{x}) = \max\left\{\|\mathbf{x}\|_2 - \frac{1}{\beta^*}, 0\right\}\frac{\mathbf{x}}{\|\mathbf{x}\|_2} := s^*$ and $\lim_{l\to+\infty}h^l(\mathbf{x}) = \mathbf{D}^*(\mathbf{M}^*)^{-1}\left[(\mathbf{D}^*)^T(\mathbf{x}) + \frac{\mu}{\beta^*}\mathbf{K}^T\mathbf{y}\right] := h^*(\mathbf{x})$ Invoking Lemma 1 and 2, we have $\|s^l(h^l(\mathbf{w}^l)) - s^l(h^l(\tilde{\mathbf{w}}))\| \le \|h^l(\mathbf{w}^l) - h^l(\tilde{\mathbf{w}})\| \le \|\mathbf{w}^l - \tilde{\mathbf{w}}\|$. Furthermore, Lemma 1 and 2 ensures non-expansiveness of the shrinkage operator s^* and the function h^*; suppose $\tilde{\mathbf{w}}$ is any fixed point of $s^*\circ h^*$, that $\tilde{\mathbf{w}} = s^*(h^*(\tilde{\mathbf{w}}))$, then we have

$$
\begin{aligned}
\|\mathbf{w}^{l+1} - \tilde{\mathbf{w}}\| &= \|s^l(h^l(\mathbf{w}^l)) - s^*(h^*(\tilde{\mathbf{w}}))\| \\
&= \|s^l(h^l(\mathbf{w}^l)) - s^l(h^l(\tilde{\mathbf{w}})) + s^l(h^l(\tilde{\mathbf{w}})) - s^*(h^*(\tilde{\mathbf{w}}))\| \\
&\le \|s^l(h^l(\mathbf{w}^l)) - s^l(h^l(\tilde{\mathbf{w}}))\| + \|s^l(h^l(\tilde{\mathbf{w}})) - s^*(h^*(\tilde{\mathbf{w}}))\| \\
&= \|\mathbf{w}^l - \tilde{\mathbf{w}}\| + \zeta^l,
\end{aligned}
\tag{8}
$$

where we write $\zeta^l = \|s^l(h^l(\tilde{\mathbf{w}})) - s^*(h^*(\tilde{\mathbf{w}}))\|$. We are going to show that $\sum_{l=1}^{l}\zeta^l$ converges, which indicates that the real non-negative sequence $\|\mathbf{w}^{l+1} - \tilde{\mathbf{w}}\|$ converges as well according to Lemma 3. The assertion holds trivially if $\|h^*(\widetilde{\mathbf{w}})\| < \frac{1}{\beta^*}$, as it will be the case that $\|D^l\widetilde{\mathbf{u}}\| < \frac{1}{\beta^l}$ for l sufficiently large. Otherwise, $\|\mathbf{D}^*\widetilde{\mathbf{u}}\| > 0$. Choose l sufficiently large so that $|\beta^l - \beta^*| \le \frac{\beta^*}{2}$, $\|(\mathbf{D}^l - \mathbf{D}^*)\widetilde{\mathbf{u}}\| < \frac{\|\mathbf{D}^*\widetilde{u}\|}{2}$. we have

$$
\begin{aligned}
\|s^l(h^l(\tilde{\mathbf{w}})) - s^*(h^*(\tilde{\mathbf{w}}))\| &= \|h^l(\tilde{\mathbf{w}}) - h^*(\tilde{\mathbf{w}}) - (\chi^l(h^l(\tilde{\mathbf{w}})) - \chi^*(h^*(\tilde{\mathbf{w}}))\| \\
&= \left\|(\mathbf{D}^l - \mathbf{D}^*)\tilde{u} - \left(\frac{1}{\beta^l}\frac{\mathbf{D}^l\tilde{u}}{\|\mathbf{D}^l\tilde{u}\|} - \frac{1}{\beta^*}\frac{\mathbf{D}^*\tilde{u}}{\|\mathbf{D}^*\tilde{u}\|}\right)\right\| \\
&\le \left\|(\mathbf{D}^l - \mathbf{D}^*)\tilde{u}\right\| + \left\|\frac{1}{\beta^l}\frac{\mathbf{D}^l\tilde{u}}{\|\mathbf{D}^l\tilde{u}\|} - \frac{1}{\beta^*}\frac{\mathbf{D}^*\tilde{u}}{\|\mathbf{D}^*\tilde{u}\|}\right\| \\
&\le \|\mathbf{D}^l - \mathbf{D}^*\|\|\widetilde{\mathbf{u}}\| + \left\|\left(\frac{1}{\beta^l} - \frac{1}{\beta^*}\right)\frac{\mathbf{D}^l\widetilde{\mathbf{u}}}{\|\mathbf{D}^l\widetilde{\mathbf{u}}\|}\right\| \\
&+ \frac{1}{\beta^*}\left\|\frac{\mathbf{D}^l\widetilde{\mathbf{u}}}{\|\mathbf{D}^l\widetilde{\mathbf{u}}\|} - \frac{\mathbf{D}^*\widetilde{\mathbf{u}}}{\|\mathbf{D}^*\widetilde{\mathbf{u}}\|}\right\| \\
&= \|\xi_l\|\|\widetilde{\mathbf{u}}\| + \frac{|\beta^l - \beta^*|}{\beta^l\beta^*} + \frac{1}{\beta^*}\left\|\frac{\|\mathbf{D}^*\widetilde{\mathbf{u}}\|\mathbf{D}^l\widetilde{\mathbf{u}} - \|\mathbf{D}^l\widetilde{\mathbf{u}}\|\mathbf{D}^*\widetilde{\mathbf{u}}}{\|\mathbf{D}^l\widetilde{\mathbf{u}}\|\|\mathbf{D}^*\widetilde{\mathbf{u}}\|}\right\| \\
&\le \|\xi_l\|\|\widetilde{\mathbf{u}}\| + \frac{2|\beta^l - \beta^*|}{{\beta^*}^2} \\
&+ \frac{2}{\beta^*}\frac{\|\|\mathbf{D}^*\widetilde{\mathbf{u}}\|(\mathbf{D}^l - \mathbf{D}^*)\widetilde{\mathbf{u}} + (\|\mathbf{D}^*\widetilde{\mathbf{u}}\| - \|\mathbf{D}^l\widetilde{\mathbf{u}}\|)\mathbf{D}^*\widetilde{\mathbf{u}}\|}{\|\mathbf{D}^*\widetilde{\mathbf{u}}\|^2} \\
&\le \|\xi_l\|\|\widetilde{\mathbf{u}}\| + \frac{2|\beta^l - \beta^*|}{{\beta^*}^2} \\
&+ \frac{2}{\beta^*}\frac{\|\mathbf{D}^*\widetilde{\mathbf{u}}\|\|(\mathbf{D}^l - \mathbf{D}^*)\widetilde{\mathbf{u}}\| + \|\mathbf{D}^*\widetilde{\mathbf{u}} - \mathbf{D}^l\widetilde{\mathbf{u}}\|\|\mathbf{D}^*\widetilde{\mathbf{u}}\|}{\|\mathbf{D}^*\widetilde{\mathbf{u}}\|^2} \\
&\le \|\xi_l\|\|\widetilde{\mathbf{u}}\| + \frac{2|\gamma_l|}{{\beta^*}^2} + \frac{2}{\beta^*}\frac{\|\xi_l\|\|\widetilde{\mathbf{u}}\|}{\|\mathbf{D}^*\widetilde{\mathbf{u}}\|}
\end{aligned}
\tag{9}
$$

where $\tilde{u} = \left({\mathbf{D}^*}^{\mathrm{T}}\mathbf{D}^* + \frac{\mu}{\beta^*}\mathbf{K}^{\mathrm{T}}\mathbf{K}\right)^{-1}\left({\mathbf{D}^*}^{\mathrm{T}}\tilde{\mathbf{w}} + \frac{\mu}{\beta^*}\mathbf{K}^{\mathrm{T}}\mathbf{y}\right)$, which is bounded. The convergence of $\sum_l\zeta^l$ follows from that of $\sum_l\|\xi_l\|$ and $\sum_l|\gamma_l|$. Now that $\{\mathbf{w}^l\}_l$ is bounded, let $\mathbf{w}^*$ be any limit point of it and $\{\mathbf{w}^{l_i}\}_i$ be a subsequence such that $\lim_{i\to\infty}\mathbf{w}^{l_i}=\mathbf{w}^*$. Let $l\to+\infty$, (8) becomes

$$\lim_{l\to+\infty}\|\mathbf{w}^{l+1}-\tilde{\mathbf{w}}\|\leq\lim_{l\to+\infty}\|\mathbf{w}^{l}-\tilde{\mathbf{w}}\|=\lim_{i\to\infty}\|\mathbf{w}^{l_i}-\tilde{\mathbf{w}}\|=\|\mathbf{w}^*-\tilde{\mathbf{w}}\| \tag{10}$$

which implies that all limit point of $\{\mathbf{w}^k\}$, if more than one, have an equal distance to $\tilde{\mathbf{w}}$. On the other hand,

$$\begin{aligned}\|\mathbf{w}^{l_i+1}-s^*(h^*(\mathbf{w}^*))\| &= \|s^{l_i}(h^{l_i}(\mathbf{w}^{l_i}))-s^*(h^*(\mathbf{w}^*))\| \\ &\leq \|s^{l_i}(h^{l_i}(\mathbf{w}^{l_i}))-s^{l_i}(h^{l_i}(\mathbf{w}^*))\| \\ &\quad+\|s^{l_i}(h^{l_i}(\mathbf{w}^*))-s^*(h^*(\mathbf{w}^*))\| \\ &\leq \|\mathbf{w}^{l_i}-\mathbf{w}^*\|+\|s^{l_i}(h^{l_i}(\mathbf{w}^*))-s^*(h^*(\mathbf{w}^*))\|,\end{aligned} \tag{11}$$

which implies $\lim_{i\to\infty}\mathbf{w}^{l_i+1}=s^*(h^*(\mathbf{w}^*))$, so that $s^*(h^*(\mathbf{w}^*))$ is also a limit point of $\{\mathbf{w}^k\}$ which is required to have the same distance to $\tilde{\mathbf{w}}$ as $\mathbf{w}^*$ does, that is $\|\mathbf{w}^*-\tilde{\mathbf{w}}\|=\|s^*(h^*(\mathbf{w}^*))-\tilde{\mathbf{w}}\|=\|s^*(h^*(\mathbf{w}^*))-s^*(h^*(\tilde{\mathbf{w}}))\|$. Since $\tilde{\mathbf{w}}$ is any fixed point of $s^*(h^*(\cdot))$, replacing $\tilde{\mathbf{w}}$ with $\mathbf{w}^*$ in (10) gives rise to $\lim_{l\to\infty}\|\mathbf{w}^l-\mathbf{w}^*\|=\lim_{i\to\infty}\|\mathbf{w}^{l_i}-\mathbf{w}^*\|=\|\mathbf{w}^*-\mathbf{w}^*\|=0$, which implies

$$\lim_{l\to\infty}\mathbf{w}^l=\mathbf{w}^* \tag{12}$$

combining (6) (in the original manuscript) and (12) leads to

$$\lim_{l\to+\infty}\mathbf{u}^l=\lim_{l\to+\infty}\left(\mathbf{D}^{l\,\mathrm{T}}\mathbf{D}^l+\frac{\mu}{\beta^l}\mathbf{K}^{\mathrm{T}}\mathbf{K}\right)^{-1}\left(\mathbf{D}^{l\,\mathrm{T}}\mathbf{w}^l+\frac{\mu}{\beta^l}\mathbf{K}^{\mathrm{T}}\mathbf{y}\right)=\mathbf{u}^*$$

which completes the proof of Theorem 1. □

[R1] Y. Wang, J. Yang, W. Yin, and Y. Zhang, “A new alternating minimization algorithm for total variation image reconstruction,” *SIAM J. Imaging Sci*, vol. 1, no. 3, pp. 248–272, 2008.