Supplementary for Locally Regularized Sparse Graph by Fast Proximal Gradient Descent

Dongfang Sun, Yingzhen Yang[1]

[1]School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85281, USA ,
{dsun30,yingzhen.yang}@asu.edu

1 PROOFS AND MORE TECHNICAL RESULTS

Proposition 1.1. *Define $\mathcal{C}^+ = \{t \colon 1 \le t \le n, c_{ti} > 0\}$, and $\mathcal{C}^- = \{t \colon 1 \le t \le n, c_{ti} < 0\}$. Let $\mathbf{z}^*$ be a critical point of function $\tilde{F}$ in eq.(7) of the main paper. Then for arbitrary small positive number $\varepsilon > 0$, $\tilde{\mathbf{z}}^{*,\epsilon} \in \mathbb{R}^n$ defined by*

$$\tilde{\mathbf{z}}_k^{*,\epsilon} = \begin{cases} \mathbf{z}_k^* & \text{if } \mathbf{z}_k^* \neq 0 \text{ or } k \in \mathcal{C}^+ \\ \varepsilon & \text{otherwise} \end{cases} \tag{1}$$

Then there exists $\mathbf{u} \in \tilde{\partial} F(\tilde{\mathbf{z}}_k^{*,\epsilon})$ for F in eq.(6) of the main paper such that $\|\mathbf{u}\|_2 \le L_f |\mathcal{C}| \varepsilon$ where $L_f \coloneqq 2\sigma_{\max}(\mathbf{X}^\top \mathbf{X})$.

Proof. Since the only different elements between $\tilde{\mathbf{z}}^{*,\epsilon}$ and $\mathbf{z}^*$ are those with indices in $\mathcal{A} = \mathcal{C}^{-1} \bigcap \{k \colon \mathbf{z}_k^* = 0\}$, we have

$$\|\nabla f(\tilde{\mathbf{z}}^*) - \nabla f(\mathbf{z}^*)\|_2 \le L_f \|\tilde{\mathbf{z}}^* - \mathbf{z}^*\|_2 \le L_f |\mathcal{C}^{-1}| \varepsilon,$$

where $L_f = 2\sigma_{\max}(\mathbf{X}^\top \mathbf{X})$. Because $\mathbf{z}^*$ be a critical point of function $\tilde{F}$, there exists $\mathbf{q} \in \tilde{\partial} h_{\gamma,c}$ such that $\mathbf{p} \coloneqq \nabla f(\mathbf{z}^*) + \mathbf{q} = \mathbf{0}$. Define $\tilde{h}_{\gamma,c} = \gamma \sum\limits_{k=1}^{n} c_{ki} \mathbb{I}_{\mathbf{Z}_k^i \neq 0}$. With the definition of $\tilde{\mathbf{z}}^{*,\epsilon}$, we have $\tilde{\mathbf{q}} \in \tilde{\partial} \tilde{h}_{\gamma,c}(\tilde{\mathbf{z}}^{*,\epsilon})$ such that $\tilde{\mathbf{q}}_k = 0$ for $k \in \mathcal{A}$ and $\tilde{\mathbf{q}}_k = \mathbf{q}_k$ otherwise. Moreover, $\mathbf{q}_k = 0$ for all $k \in \mathcal{A}$.
Therefore, let $\tilde{\mathbf{p}} \triangleq \nabla f(\tilde{\mathbf{z}}^{*,\epsilon}) + \tilde{\mathbf{q}} \in \partial F(\tilde{\mathbf{z}}^{*,\epsilon})$, we have

$$\|\tilde{\mathbf{p}}\|_2 = \|\tilde{\mathbf{p}} - \mathbf{p}\|_2 = \|\nabla f(\tilde{\mathbf{z}}^*) - \nabla f(\mathbf{z}^*)\|_2 \le L_f |\mathcal{C}^{-1}| \varepsilon.$$

The claim of this proposition follows with $\mathbf{u} = \tilde{\mathbf{p}}$.

□

We repeat critical equations in the main paper and define more notations before stating the proof of Theorem 3.2.

$$\operatorname{prox}_{sh_{\gamma,c}}(\mathbf{u}) \coloneqq \underset{\mathbf{v} \in \mathbb{R}^n, \mathbf{v}_i = 0}{\arg\min} \frac{1}{2s} \|\mathbf{v} - \mathbf{u}\|_2^2 + h_{\gamma,c}(\mathbf{z}) = T_{s,\gamma,c}(\mathbf{u}),$$

where $s > 0$ is the step size, $T_{s,\gamma,c}$ is an element-wise hard thresholding operator. For $1 \le t \le n$,

$$[T_{s,\gamma,c}(\mathbf{u})]_t = \left\{ \begin{array}{ccl} 0 & : & |\mathbf{u}_t| \le \sqrt{2s\gamma c_{ti}} \text{ and } c_{ti} > 0, \text{ or } t = i \\ \mathbf{u}_t & : & \text{otherwise} \end{array} \right. \tag{2}$$

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

1.1 PROOF OF THEOREM 3.2

Proof of Theorem 3.2. First of all, it can be verified that $\mathrm{supp}({\mathbf{z}_{\mathcal{C}}}^{(k)}) \subseteq \mathrm{supp}({\mathbf{z}_{\mathcal{C}}}^{(k-1)})$ for all $k \ge 1$ when $s < \frac{2\tau}{G^2}$. Therefore, there exists a finite $k' \ge 1$ such that $\left\{{\mathbf{z}_{\mathcal{C}}}^{(k)}\right\}_{k \ge k'}$ have the same support $\mathcal{S}$. We note that λ can be also be slightly adjusted so that $\mathrm{supp}(\mathbf{v}_{\mathcal{C}}^{(k)}) = \mathcal{S}$ for all $k \ge k_0$. Now we consider any $k > k'$ in the sequel, and let $\mathbf{z} \in \mathbb{R}^n$ be a vector such that $\mathrm{supp}(\mathbf{z}_{\mathcal{C}}) = \mathcal{S}$.

Because f have L_f-Lipschitz continuous gradient, we have

$$f(\mathbf{z}^{(k)}) \le f(\mathbf{m}^{(k)}) + \langle \nabla f(\mathbf{m}^{(k)}), \mathbf{z}^{(k)} - \mathbf{m}^{(k)} \rangle + \frac{L_f}{2} \left\| \mathbf{z}^{(k)} - \mathbf{m}^{(k)} \right\|_2^2. \tag{3}$$

Also,

$$\begin{aligned}
& f(\mathbf{m}^{(k)}) - (1-\alpha_k) f(\mathbf{z}^{(k-1)}) - \alpha_k f(\mathbf{z}) \\
& = (1-\alpha_k) \left(f(\mathbf{m}^{(k)}) - f(\mathbf{z}^{(k-1)}) \right) + \alpha_k \left(f(\mathbf{m}^{(k)}) - f(\mathbf{z}) \right) \\
& \overset{①}{\le} (1-\alpha_k) \langle \nabla f(\mathbf{m}^{(k)}), \mathbf{m}^{(k)} - \mathbf{z}^{(k-1)} \rangle + \alpha_k \langle \nabla f(\mathbf{m}^{(k)}), \mathbf{m}^{(k)} - \mathbf{z} \rangle \\
& \le \langle \nabla f(\mathbf{m}^{(k)}), (1-\alpha_k)(\mathbf{m}^{(k)} - \mathbf{z}^{(k-1)}) + \alpha_k (\mathbf{m}^{(k)} - \mathbf{z}) \rangle \\
& = \langle \nabla f(\mathbf{m}^{(k)}), \mathbf{m}^{(k)} - (1-\alpha_k)\mathbf{z}^{(k-1)} - \alpha_k \mathbf{z} \rangle,
\end{aligned} \tag{4}$$

where ① is due to the convexity of f.

We have $\tilde{\mathbf{v}}^{(k)} = \mathbf{v}^{(k-1)} - \lambda_k \nabla f(\mathbf{m}^{(k)})$, and it follows that

$$\begin{aligned}
& \frac{1}{2\lambda_k} \left(\left\| \mathbf{v}^{(k-1)} - \mathbf{z} \right\|_2^2 - \left\| \mathbf{v}^{(k)} - \mathbf{z} \right\|_2^2 - \left\| \mathbf{v}^{(k)} - \mathbf{v}^{(k-1)} \right\|_2^2 \right) \\
& = \frac{1}{\lambda_k} \langle \mathbf{z} - \mathbf{v}^{(k)}, \mathbf{v}^{(k)} - \mathbf{v}^{(k-1)} \rangle \\
& \overset{①}{=} \frac{1}{\lambda_k} \langle \mathbf{z} - \mathbf{v}^{(k)}, \tilde{\mathbf{v}}^{(k)} - \mathbf{v}^{(k-1)} \rangle \\
& = \langle \nabla f(\mathbf{m}^{(k)}), \mathbf{v}^{(k)} - \mathbf{z} \rangle,
\end{aligned} \tag{5}$$

and ① is due to the fact that $\mathrm{supp}(\mathbf{z}_{\mathcal{C}} - \mathbf{v}_{\mathcal{C}}^{(k)}) \subseteq \mathcal{S}$ because $\mathrm{supp}(\mathbf{z}_{\mathcal{C}}) = \mathcal{S}$, $\mathrm{supp}(\mathbf{v}_{\mathcal{C}}^{(k)}) \subseteq \mathcal{S}$.

Because $\mathrm{supp}(\mathbf{v}_{\mathcal{C}}^{(k)}) \subseteq \mathrm{supp}(\mathbf{z}_{\mathcal{C}})$, we have

$$h_{\gamma,c}(\mathbf{v}^{(k)}) \le h_{\gamma,c}(\mathbf{z}). \tag{6}$$

It follows by (5) and (6) that

$$\begin{aligned}
& \langle \nabla f(\mathbf{m}^{(k)}), \mathbf{v}^{(k)} - \mathbf{z} \rangle + h_{\gamma,c}(\mathbf{v}^{(k)}) \\
& \le h_{\gamma,c}(\mathbf{z}) + \frac{1}{2\lambda_k} \left(\left\| \mathbf{v}^{(k-1)} - \mathbf{z} \right\|_2^2 - \left\| \mathbf{v}^{(k)} - \mathbf{z} \right\|_2^2 - \left\| \mathbf{v}^{(k)} - \mathbf{v}^{(k-1)} \right\|_2^2 \right)
\end{aligned} \tag{7}$$

Similar to (5), we have

$$\frac{1}{2s} \left(\left\| \mathbf{m}^{(k)} - \mathbf{z} \right\|_2^2 - \left\| \mathbf{z}^{(k)} - \mathbf{z} \right\|_2^2 - \left\| \mathbf{z}^{(k)} - \mathbf{m}^{(k)} \right\|_2^2 \right) = \frac{1}{s} \langle \mathbf{z} - \mathbf{z}^{(k)}, \mathbf{z}^{(k)} - \mathbf{m}^{(k)} \rangle. \tag{8}$$

For any $\mathbf{q} \in \partial h_{\gamma,c}(\mathbf{z}^{(k)})$, due to the fact that $\mathrm{supp}(\mathbf{z}_{\mathcal{C}}) = \mathrm{supp}(\mathbf{z}_{\mathcal{C}}^{(k)})$,

$$\langle \mathbf{z} - \mathbf{z}^{(k)}, \mathbf{q} \rangle + h_{\gamma,c}(\mathbf{z}^{(k)}) = h_{\gamma,c}(\mathbf{z}). \tag{9}$$

By (8) and (9),

$$\begin{aligned}&\langle \mathbf{z}-\mathbf{z}^{(k)}, \frac{1}{s}(\mathbf{z}^{(k)}-\mathbf{m}^{(k)})+\mathbf{q}\rangle + h_{\gamma,c}(\mathbf{z}^{(k)})\\ &= h_{\gamma,c}(\mathbf{z}) + \frac{1}{2s}\left(\left\|\mathbf{m}^{(k)}-\mathbf{z}\right\|_2^2 - \left\|\mathbf{z}^{(k)}-\mathbf{z}\right\|_2^2 - \left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2\right)\end{aligned} \tag{10}$$

By the optimality condition of the proximal mapping in eq.(10) in Algorithm 1, we can choose $\mathbf{q}\in\partial h_{\gamma,c}(\mathbf{z}^{(k)})$ such that $\mathbf{z}^{(k)}=\mathbf{m}^{(k)}-s\left(\nabla f(\mathbf{m}^{(k)})+\mathbf{q}\right)$. Plugging such $\mathbf{q}$ in (10), we have

$$\langle \nabla f(\mathbf{m}^{(k)}), \mathbf{z}^{(k)}-\mathbf{z}\rangle + h_{\gamma,c}(\mathbf{z}^{(k)}) = h_{\gamma,c}(\mathbf{z}) + \frac{1}{2s}\left(\left\|\mathbf{m}^{(k)}-\mathbf{z}\right\|_2^2 - \left\|\mathbf{z}^{(k)}-\mathbf{z}\right\|_2^2 - \left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2\right) \tag{11}$$

Setting $\mathbf{z}=(1-\alpha_k)\mathbf{z}^{(k-1)}+\alpha_k\mathbf{v}^{(k)}$ in (11), we have

$$\begin{aligned}&\langle \nabla f(\mathbf{m}^{(k)}), \mathbf{z}^{(k)}-(1-\alpha_k)\mathbf{z}^{(k-1)}-\alpha_k\mathbf{v}^{(k)}\rangle + h_{\gamma,c}(\mathbf{z}^{(k)})\\ &\le h_{\gamma,c}((1-\alpha_k)\mathbf{z}^{(k-1)}+\alpha_k\mathbf{v}^{(k)}) + \frac{1}{2s}\left(\left\|\mathbf{m}^{(k)}-(1-\alpha_k)\mathbf{z}^{(k-1)}-\alpha_k\mathbf{v}^{(k)}\right\|_2^2 - \left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2\right)\\ &\overset{①}{\le} (1-\alpha_k)h_{\gamma,c}(\mathbf{z}^{(k-1)})+\alpha_k h_{\gamma,c}(\mathbf{v}^{(k)})\\ &\quad+\frac{1}{2s}\left(\left\|\mathbf{m}^{(k)}-(1-\alpha_k)\mathbf{z}^{(k-1)}-\alpha_k\mathbf{v}^{(k)}\right\|_2^2 - \left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2\right)\\ &\overset{②}{\le} (1-\alpha_k)h_{\gamma,c}(\mathbf{z}^{(k-1)})+\alpha_k h_{\gamma,c}(\mathbf{v}^{(k)}) + \frac{1}{2s}\left(\alpha_k^2\left\|\mathbf{v}^{(k)}-\mathbf{v}^{(k-1)}\right\|_2^2 - \left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2\right),\end{aligned} \tag{12}$$

where ① is due to the fact that $\mathrm{supp}(\mathbf{v}_{\mathcal{C}}^{(k)})=\mathrm{supp}(\mathbf{z}_{\mathcal{C}}^{(k-1)}))$ and $h_{\gamma,c}$ satisfies $h_{\gamma,c}\left((1-\tau)\mathbf{u}+\tau\mathbf{v}\right)\le(1-\tau)h_{\gamma,c}(\mathbf{u})+\tau h_{\gamma,c}(\mathbf{v})$ for any two vectors $\mathbf{u}$, $\mathbf{v}$ with $\mathrm{supp}(\mathbf{u}_{\mathcal{C}})=\mathrm{supp}(\mathbf{v}_{\mathcal{C}})$ and any $\tau\in(0,1)$. ② is due to the fact that $\mathbf{m}^{(k)}-(1-\alpha_k)\mathbf{z}^{(k-1)}-\alpha_k\mathbf{v}^{(k)}=\alpha_k(\mathbf{v}^{(k-1)}-\mathbf{v}^{(k)})$ according to eq.(9) in Algorithm 1.

Computing $\alpha_k\times$ (7) + (12), we have

$$\begin{aligned}&\langle \nabla f(\mathbf{m}^{(k)}), \mathbf{z}^{(k)}-(1-\alpha_k)\mathbf{z}^{(k-1)}-\alpha_k\mathbf{z}\rangle + h_{\gamma,c}(\mathbf{z}^{(k)})\\ &\le (1-\alpha_k)h_{\gamma,c}(\mathbf{z}^{(k-1)})+\alpha_k h_{\gamma,c}(\mathbf{z})\\ &\quad+\frac{\alpha_k}{2\lambda_k}\left(\left\|\mathbf{v}^{(k-1)}-\mathbf{z}\right\|_2^2-\left\|\mathbf{v}^{(k)}-\mathbf{z}\right\|_2^2\right)+\left(\frac{\alpha_k^2}{2s}-\frac{\alpha_k}{2\lambda_k}\right)\left\|\mathbf{v}^{(k)}-\mathbf{v}^{(k-1)}\right\|_2^2-\frac{1}{2s}\left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2\\ &\overset{①}{\le} (1-\alpha_k)h_{\gamma,c}(\mathbf{z}^{(k-1)})+\alpha_k h_{\gamma,c}(\mathbf{z})+\frac{\alpha_k}{2\lambda_k}\left(\left\|\mathbf{v}^{(k-1)}-\mathbf{z}\right\|_2^2-\left\|\mathbf{v}^{(k)}-\mathbf{z}\right\|_2^2\right)-\frac{1}{2s}\left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2,\end{aligned} \tag{13}$$

and ① is due to $\lambda_k\alpha_k\le s$.

Combining (3), (4) and (13), and noting that $\tilde{F}(\mathbf{z})=f(\mathbf{z})+h_{\gamma,c}(\mathbf{z})$, we have

$$\begin{aligned}\tilde{F}(\mathbf{z}^{(k)}) &\le (1-\alpha_k)\tilde{F}(\mathbf{z}^{(k-1)})+\alpha_k\tilde{F}(\mathbf{z})-\left(\frac{1}{2s}-\frac{L_f}{2}\right)\left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2\\ &\quad+\frac{\alpha_k}{2\lambda_k}\left(\left\|\mathbf{v}^{(k-1)}-\mathbf{z}\right\|_2^2-\left\|\mathbf{v}^{(k)}-\mathbf{z}\right\|_2^2\right).\end{aligned} \tag{14}$$

It follows by (14) that

$$\begin{aligned}\tilde{F}(\mathbf{z}^{(k)})-\tilde{F}(\mathbf{z}) &\le (1-\alpha_k)\left(\tilde{F}(\mathbf{z}^{(k-1)})-\tilde{F}(\mathbf{z})\right)\\ &\quad-\left(\frac{1}{2s}-\frac{L_f}{2}\right)\left\|\mathbf{z}^{(k)}-\mathbf{m}^{(k)}\right\|_2^2+\frac{\alpha_k}{2\lambda_k}\left(\left\|\mathbf{v}^{(k-1)}-\mathbf{z}\right\|_2^2-\left\|\mathbf{v}^{(k)}-\mathbf{z}\right\|_2^2\right).\end{aligned} \tag{15}$$

Define a sequence $\{T_k\}_{k=1}^{\infty}$ as $T_1 = 1$, and $T_k = (1-\alpha_k)T_{k-1}$ for $k \geq 2$. Dividing both sides of (15) by T_k, we have

$$\begin{aligned}\frac{\tilde{F}(\mathbf{z}^{(k)}) - \tilde{F}(\mathbf{z})}{T_k} &\leq \frac{\tilde{F}(\mathbf{z}^{(k-1)}) - \tilde{F}(\mathbf{z})}{T_{k-1}} - \frac{1-L_f s}{2sT_k}\left\|\mathbf{z}^{(k)} - \mathbf{m}^{(k)}\right\|_2^2 \\ &+ \frac{\alpha_k}{2\lambda_k T_k}\left(\left\|\mathbf{v}^{(k-1)} - \mathbf{z}\right\|_2^2 - \left\|\mathbf{v}^{(k)} - \mathbf{z}\right\|_2^2\right). \end{aligned} \tag{16}$$

Since we choose $\alpha_k = \frac{2}{k+1}$, it follows that $T_k = \frac{2}{k(k+1)}$ for all $k \geq 1$. Plugging the values of α_k and T_k in $\frac{\alpha_k}{2\lambda_k T_k}$ in (16), we have

$$\begin{aligned}\frac{\tilde{F}(\mathbf{z}^{(k)}) - \tilde{F}(\mathbf{z})}{T_k} &\leq \frac{\tilde{F}(\mathbf{z}^{(k-1)}) - \tilde{F}(\mathbf{z})}{T_{k-1}} - \frac{1-L_f s}{2sT_k}\left\|\mathbf{z}^{(k)} - \mathbf{m}^{(k)}\right\|_2^2 \\ &+ \frac{k}{2\lambda_k}\left(\left\|\mathbf{v}^{(k-1)} - \mathbf{z}\right\|_2^2 - \left\|\mathbf{v}^{(k)} - \mathbf{z}\right\|_2^2\right) \\ &\overset{①}{\leq} \frac{\tilde{F}(\mathbf{z}^{(k-1)}) - \tilde{F}(\mathbf{z})}{T_{k-1}} - \frac{1-L_f s}{2sT_k}\left\|\mathbf{z}^{(k)} - \mathbf{m}^{(k)}\right\|_2^2 + \frac{k}{2\lambda_k}\left\|\mathbf{v}^{(k-1)} - \mathbf{z}\right\|_2^2 \\ &- \frac{k+1}{2\lambda_{k+1}}\left\|\mathbf{v}^{(k)} - \mathbf{z}\right\|_2^2, \end{aligned} \tag{17}$$

where ① is due to the condition that $\lambda_{k+1} \geq \frac{k+1}{k}\lambda_k$ for $k \geq 1$.

Set $k_0 = k' + 1$. Summing the above inequality for $k = k_0, \ldots, m$ with $m \geq k_0$, we have

$$\begin{aligned}\frac{\tilde{F}(\mathbf{z}^{(m)}) - \tilde{F}(\mathbf{z})}{T_m} &\leq \frac{\tilde{F}(\mathbf{z}^{(k_0-1)}) - \tilde{F}(\mathbf{z})}{T_{k_0-1}} + \frac{k_0\left\|\mathbf{v}^{(k_0-1)} - \mathbf{z}\right\|_2^2}{2\lambda_{k_0}} - \sum_{k=k_0}^{m}\frac{1-L_f s}{2sT_k}\left\|\mathbf{z}^{(k)} - \mathbf{m}^{(k)}\right\|_2^2 \\ &\leq \frac{k_0(k_0-1)\left(\tilde{F}(\mathbf{z}^{(k_0-1)}) - \tilde{F}(\mathbf{z})\right)}{2} + \frac{\left\|\mathbf{v}^{(k_0-1)} - \mathbf{z}\right\|_2^2}{2\eta}. \end{aligned} \tag{18}$$

Since $T_m = \frac{2}{m(m+1)}$, it follows by (18) with $z = z^*$ that

$$\tilde{F}(\mathbf{z}^{(m)}) - \tilde{F}(\mathbf{z}^*) \leq \frac{1}{m(m+1)} \cdot \left(k_0(k_0-1)\left(\tilde{F}(\mathbf{z}^{(k_0-1)}) - \tilde{F}(\mathbf{z}^*)\right) + \frac{\left\|\mathbf{v}^{(k_0-1)} - \mathbf{z}^*\right\|_2^2}{\eta}\right). \tag{19}$$

Changing m to k in (19) completes the proof.

□

2 ADDITIONAL ILLUSTRATION

Figure 1 illustrates the comparison between the weighed adjacency matrix of ℓ^1-graph and SRSG.

References

D.J. Newman A. Asuncion. UCI machine learning repository, 2007.

D. Plummer and L. Lovász. *Matching Theory*. North-Holland Mathematics Studies. Elsevier Science, 1986.

Xin Zheng, Deng Cai, Xiaofei He, Wei-Ying Ma, and Xueyin Lin. Locality preserving clustering for image database. In *Proceedings of the 12th Annual ACM International Conference on Multimedia*, MULTIMEDIA '04, pages 885–891, New York, NY, USA, 2004. ACM.

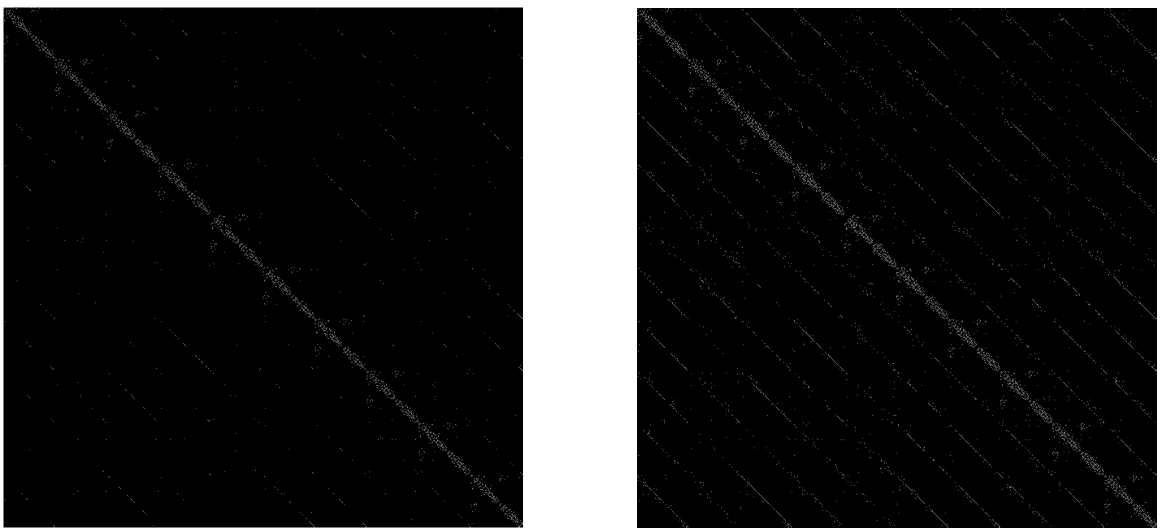

Figure 1: The comparison between the weighed adjacency matrix W of the sparse graph produced by ℓ^1-graph (right) and SRSG (left) on the Extended Yale Face Database B, where each white dot indicates an edge in the sparse graph.