Appendix of "Adaptive Consensus Clustering for Multiple K-means via Base Results Refining"

Appendix A: Derivation of Eq.(10)

We have

$$\begin{aligned}\min_{\mathbf{A}} \mathcal{J}_1 = & vtr\left(\mathbf{A}^k\mathbf{X}\mathbf{X}^T\mathbf{A}^{kT}\right) - 2tr\left(\sum_{m=1}^{v}\mathbf{G}^{(m)}\mathbf{F}^{(m)}\mathbf{X}^T\mathbf{A}^k\right) \\ & + tr(\mathbf{\Lambda}_1^T\mathbf{A}) + \frac{\mu}{2}\|\mathbf{A} - (\mathbf{I} + \mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}})/2\|_F^2. \end{aligned} \tag{1}$$

Let $\mathcal{L}_1 = tr\left(\mathbf{A}^k\mathbf{X}\mathbf{X}^T\mathbf{A}^{kT}\right)$, $\mathcal{L}_2 = tr\left(\sum_{m=1}^{v}\mathbf{G}^{(m)}\mathbf{F}^{(m)}\mathbf{X}^T\mathbf{A}^k\right)$, $\mathcal{L}_3 = tr(\mathbf{\Lambda}_1^T\mathbf{A})$, and $\mathcal{L}_4 = \|\mathbf{A} - (\mathbf{I} + \mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}})/2\|_F^2$. Then we compute the partial derivative of $\mathcal{L}_1$, $\mathcal{L}_2$, $\mathcal{L}_3$, and $\mathcal{L}_4$ w.r.t. $\mathbf{A}$ respectively.

Defining an auxiliary variable $\mathbf{U} = \mathbf{A}^k$, according to the chain rule, we have

$$\begin{aligned}\frac{\partial \mathcal{L}_1}{\partial A_{ij}} = & tr\left(\left(\frac{\partial tr(\mathbf{U}\mathbf{X}\mathbf{X}^T\mathbf{U}^T)}{\partial \mathbf{U}}\right)^T \frac{\partial \mathbf{U}}{\partial A_{ij}}\right) \\ = & 2tr\left(\mathbf{X}\mathbf{X}^T\mathbf{U}^T\frac{\partial \mathbf{A}^k}{\partial A_{ij}}\right) \\ = & 2tr\left(\mathbf{X}\mathbf{X}^T\mathbf{U}^T\sum_{r=0}^{k-1}\mathbf{A}^r\mathbf{J}^{ij}\mathbf{A}^{k-1-r}\right) \\ = & 2\sum_{r=0}^{k-1} tr\left(\mathbf{A}^{k-1-r}\mathbf{X}\mathbf{X}^T(\mathbf{A}^k)^T\mathbf{A}^r\mathbf{J}^{ij}\right) \end{aligned} \tag{2}$$

where $\mathbf{J}^{ij} \in \mathbb{R}^{n\times n}$ is a single-entry matrix, whose the (i,j)-th element is 1 and other elements are all 0s. Then, it is easy to verify that

$$\frac{\partial \mathcal{L}_1}{\partial \mathbf{A}} = 2\sum_{r=0}^{k-1}\left(\mathbf{A}^{k-1-r}\mathbf{X}\mathbf{X}^T(\mathbf{A}^k)^T\mathbf{A}^r\right)^T. \tag{3}$$

Denoting $\mathbf{B} = \sum_{m=1}^{v}\mathbf{X}\mathbf{F}^{(m)T}\mathbf{G}^{(m)T}$, we have

$$\frac{\partial \mathcal{L}_2}{\partial \mathbf{A}} = \frac{\partial tr(\mathbf{B}\mathbf{A}^k)}{\partial \mathbf{A}} = \sum_{r=0}^{k-1}(\mathbf{A}^r\mathbf{B}\mathbf{A}^{k-r-1})^T. \tag{4}$$

Moreover, we have

$$\frac{\partial \mathcal{L}_3}{\partial \mathbf{A}} = \mathbf{\Lambda}_1. \tag{5}$$

and by denoting $\mathbf{C} = (\mathbf{I} + \mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}})/2$, we have

$$\frac{\partial \mathcal{L}_4}{\partial \mathbf{A}} = 2(\mathbf{A} - \mathbf{C}). \tag{6}$$

To sum up Eqs.(3), (4), (5), (6), we obtain

$$\frac{\partial \mathcal{J}_1}{\partial \mathbf{A}} = 2v \sum_{r=0}^{k-1} \left(\mathbf{A}^{k-r-1}\mathbf{X}\mathbf{X}^T\mathbf{A}^{kT}\mathbf{A}^r\right)^T - 2\sum_{r=0}^{k-1}\left(\mathbf{A}^r\mathbf{B}\mathbf{A}^{k-r-1}\right)^T + \mathbf{\Lambda}_1 + \mu(\mathbf{A} - \mathbf{C}). \tag{7}$$

Appendix B: Derivation of Eq.(12)

We have

$$\begin{aligned}\min_{\mathbf{W}} \mathcal{J}_2 = & \lambda\|\mathbf{W} - \mathbf{F}\|_F^2 - \frac{1}{2}tr(\mathbf{\Lambda}_1^T\mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}) \\ & + tr(\mathbf{\Lambda}_2^T\mathbf{W}) + \frac{\mu}{8}tr(\mathbf{D}^{-\frac{1}{2}}\mathbf{W}^T\mathbf{D}^{-1}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}) \\ & - \frac{\mu}{2}tr\left(\mathbf{E}^T\mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}\right) + \frac{\mu}{2}\|\mathbf{W} - \mathbf{V}\|_F^2,\end{aligned} \tag{8}$$

Denote $\mathcal{I}_1 = \|\mathbf{W} - \mathbf{F}\|_F^2$, $\mathcal{I}_2 = tr(\mathbf{\Lambda}_1^T\mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}})$, $\mathcal{I}_3 = tr(\mathbf{\Lambda}_2^T\mathbf{W})$, $\mathcal{I}_4 = tr(\mathbf{D}^{-\frac{1}{2}}\mathbf{W}^T\mathbf{D}^{-1}\mathbf{W}\mathbf{D}^{-\frac{1}{2}})$, $\mathcal{I}_5 = tr\left(\mathbf{E}\mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}\right)$, and $\mathcal{I}_6 = \|\mathbf{W} - \mathbf{V}\|_F^2$.

We have

$$\frac{\partial \mathcal{I}_1}{\partial \mathbf{W}} = 2(\mathbf{W} - \mathbf{F}), \tag{9}$$

$$\frac{\partial \mathcal{I}_3}{\partial \mathbf{W}} = \mathbf{\Lambda}_2, \tag{10}$$

and

$$\frac{\partial \mathcal{I}_6}{\partial \mathbf{W}} = 2(\mathbf{W} - \mathbf{V}). \tag{11}$$

Now, we consider the derivative of $\mathcal{I}_2$. Taking $D_{ii} = \sum_{r=1}^{n} W_{ir}$ into $\mathcal{I}_2$, we have

$$
\begin{aligned}
\frac{\partial \mathcal{I}_2}{\partial W_{pq}} =& \frac{\partial \sum_{i,j=1}^{n} \Lambda_{1ij} W_{ij} \frac{1}{\sqrt{\sum_{r=1}^{n} W_{ir}}} \frac{1}{\sqrt{\sum_{r=1}^{n} W_{jr}}}}{\partial W_{pq}} \qquad (12)\\
=& \frac{\Lambda_{1pq}}{\sqrt{\sum_{r=1}^{n} W_{pr}}\sqrt{\sum_{r=1}^{n} W_{qr}}} - \frac{1}{2}\sum_{j=1}^{n} \frac{\Lambda_{1pj} W_{pj}}{\left(\sqrt{\sum_{r=1}^{n} W_{pr}}\right)^3 \sqrt{\sum_{r=1}^{n} W_{jr}}}\\
&- \frac{1}{2}\sum_{i=1}^{n} \frac{\Lambda_{1ip} W_{ip}}{\left(\sqrt{\sum_{r=1}^{n} W_{pr}}\right)^3 \sqrt{\sum_{r=1}^{n} W_{ir}}}\\
=& \frac{\Lambda_{1pq}}{\sqrt{\sum_{r=1}^{n} W_{pr}}\sqrt{\sum_{r=1}^{n} W_{qr}}} - \frac{1}{2}\frac{1}{\left(\sqrt{\sum_{r=1}^{n} W_{pr}}\right)^3} \sum_{i=1}^{n} \frac{\Lambda_{1pi} W_{pi} + \Lambda_{1ip} W_{ip}}{\sqrt{\sum_{r=1}^{n} W_{ir}}}
\end{aligned}
$$

Reformulating it to the matrix form, we have

$$
\frac{\partial \mathcal{I}_2}{\partial \mathbf{W}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{\Lambda}_1 \mathbf{D}^{-\frac{1}{2}} - \frac{1}{2} diag\left(\mathbf{D}^{-\frac{3}{2}}\left(\mathbf{\Lambda}_1^T \mathbf{D}^{-\frac{1}{2}} \mathbf{W} + \mathbf{\Lambda}_1 \mathbf{D}^{-\frac{1}{2}} \mathbf{W}^T\right)\right) \mathbf{1}^T. \qquad (13)
$$

Similarly, the derivative of $\mathbf{I}_5$ is

$$
\frac{\partial \mathcal{I}_5}{\partial \mathbf{W}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{E} \mathbf{D}^{-\frac{1}{2}} - \frac{1}{2} diag\left(\mathbf{D}^{-\frac{3}{2}}\left(\mathbf{E}^T \mathbf{D}^{-\frac{1}{2}} \mathbf{W} + \mathbf{E} \mathbf{D}^{-\frac{1}{2}} \mathbf{W}^T\right)\right) \mathbf{1}^T. \qquad (14)
$$

Now, we consider $\mathcal{I}_4$. We have

$$
\begin{aligned}
\frac{\partial \mathcal{I}_4}{\partial W_{pq}} =& \frac{\partial \left(\sum_{i,j=1}^{n} \frac{1}{\sum_{r=1}^{n} W_{ir}} \frac{1}{\sum_{r=1}^{n} W_{jr}} W_{ji}^2\right)}{\partial W_{pq}} \qquad (15)\\
=& -\sum_{j=1}^{n} \frac{W_{jp}^2}{\sum_{r=1}^{n} W_{jr}} \frac{1}{(\sum_{r=1}^{n} W_{pr})^2} - \sum_{i=1}^{n} \frac{W_{pi}^2}{\sum_{r=1}^{n} W_{ir}} \frac{1}{(\sum_{r=1}^{n} W_{pr})^2} + \frac{2W_{pq}}{\sum_{r=1}^{n} W_{pr} \sum_{r=1}^{n} W_{qr}}\\
=& \frac{2W_{pq}}{\sum_{r=1}^{n} W_{pr} \sum_{r=1}^{n} W_{qr}} - \frac{1}{(\sum_{r=1}^{n} W_{pr})^2} \sum_{i=1}^{n} \frac{W_{ip}^2 + W_{pi}^2}{\sum_{r=1}^{n} W_{ir}}.
\end{aligned}
$$

Reformulating it to the matrix form, we have

$$
\frac{\partial \mathcal{I}_4}{\partial \mathbf{W}} = 2\mathbf{D}^{-1} \mathbf{W} \mathbf{D}^{-1} - diag\left(\mathbf{D}^{-2}\left(\mathbf{W}^T \mathbf{D}^{-1} \mathbf{W} + \mathbf{W} \mathbf{D}^{-1} \mathbf{W}^T\right)\right) \mathbf{1}^T. \qquad (16)
$$

To sum up Eqs.(9), (13), (10), (16), (14) and (11), we obtain

$$
\begin{aligned}
\frac{\partial \mathcal{J}_2}{\partial \mathbf{W}} =& 2\lambda(\mathbf{W} - \mathbf{F}) - \frac{1}{2} \mathbf{D}^{-\frac{1}{2}} (\mathbf{\Lambda}_1 + \mu \mathbf{E}) \mathbf{D}^{-\frac{1}{2}}\\
&+ \frac{1}{4} diag\left(\mathbf{D}^{-\frac{3}{2}}\left((\mathbf{\Lambda_1} + \mu \mathbf{E})^T \mathbf{D}^{-\frac{1}{2}} \mathbf{W} + (\mathbf{\Lambda_1} + \mu \mathbf{E}) \mathbf{D}^{-\frac{1}{2}} \mathbf{W}^T\right)\right) \mathbf{1}^T\\
&+ \mathbf{\Lambda}_2 + \frac{\mu}{4} \mathbf{D}^{-1} \mathbf{W} \mathbf{D}^{-1} + \mu(\mathbf{W} - \mathbf{V}) \qquad (17)\\
&- \frac{\mu}{8} diag\left(\mathbf{D}^{-2}\left(\mathbf{W}^T \mathbf{D}^{-1} \mathbf{W} + \mathbf{W} \mathbf{D}^{-1} \mathbf{W}^T\right)\right) \mathbf{1}^T,
\end{aligned}
$$