A. Proofs

A.1. Proof of Theorem 2

The derivative of $g(\rho)$ is

$$g'(\rho) = \sum_{k=k_0}^{nc} \frac{2z_k^2}{(\gamma\lambda_{PQ,k} - \rho)^3}.$$

Hence, $g'(\rho) > 0$ whenever $\rho < \gamma\lambda_{PQ,k_0}$, and $g(\rho)$ is strictly increasing in the interval $(-\infty, \gamma\lambda_{PQ,k_0})$. Moreover,

$$\lim_{\rho\to-\infty} g(\rho) = -\tau^2 \quad \text{and} \quad \lim_{\rho\to\gamma\lambda_{PQ,k_0}} g(\rho) = +\infty,$$

and thus $g(\rho)$ has exactly one root in $(-\infty, \gamma\lambda_{PQ,k_0})$. Notice that $\|\boldsymbol{z}\|_2 = \|\operatorname{vec}(V_Q^\top Y V_P)\|_2 = \|V_{PQ}^\top \boldsymbol{y}\|_2 = \|\boldsymbol{y}\|_2$ since V_{PQ} is an orthonormal matrix, and then $\rho_0 = \gamma\lambda_{PQ,k_0} - \|\boldsymbol{y}\|_2/\tau = \gamma\lambda_{PQ,k_0} - \|\boldsymbol{z}\|_2/\tau$. As a result,

$$
\begin{aligned}
g(\rho_0) &= \sum_{k=k_0}^{nc} \frac{z_k^2}{(\gamma\lambda_{PQ,k} - \rho_0)^2} - \tau^2 \\
&= \sum_{k=k_0}^{nc} \frac{z_k^2}{(\gamma\lambda_{PQ,k} - \gamma\lambda_{PQ,k_0} + \|\boldsymbol{z}\|_2/\tau)^2} - \tau^2 \\
&\leq \sum_{k=k_0}^{nc} \frac{z_k^2}{(\|\boldsymbol{z}\|_2/\tau)^2} - \tau^2 \\
&= \left(\frac{\sum_{k=k_0}^{nc} z_k^2}{\|\boldsymbol{z}\|_2^2} - 1 \right)\tau^2 \\
&\leq 0,
\end{aligned}
$$

where the first inequality is because $\lambda_{PQ,k} \geq \lambda_{PQ,k_0}$ for $k \geq k_0$. The fact that $g(\rho_0) \leq 0$ concludes that the only root in $(-\infty, \gamma\lambda_{PQ,k_0})$ is in $[\rho_0, \gamma\lambda_{PQ,k_0})$ but not $(-\infty, \rho_0)$. $\qquad\square$

A.2. Proof of Theorem 3

Denote by $\boldsymbol{h} = \operatorname{vec}(H)$, $\boldsymbol{y} = \operatorname{vec}(Y)$ and $M = (\gamma P \otimes Q - \rho I_{nc})$, and denote by $\boldsymbol{h}'$, $\boldsymbol{y}'$ and M' similarly. Let $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ be two functions extracting the smallest and largest eigenvalues of a matrix. Under our assumption,

$$\lambda_{\min}(M) = \gamma\lambda_{PQ,1} - \rho \geq C_{\gamma,\tau} > 0$$

which means that M is positive definite, and so is M'. By Eq. (14),

$$
\begin{aligned}
\boldsymbol{h} - \boldsymbol{h}' &= M^{-1}\boldsymbol{y} - M'^{-1}\boldsymbol{y}' \\
&= M^{-1}(\boldsymbol{y} - \boldsymbol{y}') + (M^{-1} - M'^{-1})\boldsymbol{y}' \\
&= M^{-1}(\boldsymbol{y} - \boldsymbol{y}') + M^{-1}(M' - M)M'^{-1}\boldsymbol{y}' \\
&= M^{-1}(\boldsymbol{y} - \boldsymbol{y}') + (\rho' - \rho)M^{-1}M'^{-1}\boldsymbol{y}'.
\end{aligned}
$$

Note that $\|A\boldsymbol{v}\|_2 \leq \lambda_{\max}(A)\|\boldsymbol{v}\|_2$ for any symmetric positive-definite matrix A and any vector $\boldsymbol{v}$, as well as $\lambda_{\max}(AB) \leq \lambda_{\max}(A)\lambda_{\max}(B)$ for any symmetric positive-definite matrices A and B. Hence,

$$
\begin{aligned}
\|\boldsymbol{h} - \boldsymbol{h}'\|_2 &= \|M^{-1}(\boldsymbol{y} - \boldsymbol{y}') + (\rho' - \rho)M^{-1}M'^{-1}\boldsymbol{y}'\|_2 \\
&\leq \|M^{-1}(\boldsymbol{y} - \boldsymbol{y}')\|_2 + |\rho - \rho'|\|M^{-1}M'^{-1}\boldsymbol{y}'\|_2 \\
&\leq \lambda_{\max}(M^{-1})\|\boldsymbol{y} - \boldsymbol{y}'\|_2 + \lambda_{\max}(M^{-1})\lambda_{\max}(M'^{-1})|\rho - \rho'|\|\boldsymbol{y}'\|_2 \\
&\leq \frac{\|\boldsymbol{y} - \boldsymbol{y}'\|_2}{C_{\gamma,\tau}} + \frac{|\rho - \rho'|\|\boldsymbol{y}'\|_2}{C_{\gamma,\tau}^2},
\end{aligned}
$$

where the first inequality is the triangle inequality, the second inequality is because M^{-1} and M'^{-1} are symmetric positive definite, and the third inequality follows from $\lambda_{\max}(M^{-1}) = 1/\lambda_{\min}(M)$ and $\lambda_{\max}(M'^{-1}) = 1/\lambda_{\min}(M')$. Due to the symmetry of $\boldsymbol{h}$ and $\boldsymbol{h}'$,

$$\|\boldsymbol{h} - \boldsymbol{h}'\|_2 \leq \frac{\|\boldsymbol{y} - \boldsymbol{y}'\|_2}{C_{\gamma,\tau}} + \frac{|\rho - \rho'| \min\{\|\boldsymbol{y}\|_2, \|\boldsymbol{y}'\|_2\}}{C_{\gamma,\tau}^2}.$$

This inequality is the vectorization of (18).

For MAVR in optimization (9), Theorem 2 together with our assumption indicates that

$$\gamma\lambda_{PQ,1} - \|\boldsymbol{y}\|_2/\tau \leq \rho < \gamma\lambda_{PQ,1},$$
$$\gamma\lambda_{PQ,1} - \|\boldsymbol{y}'\|_2/\tau \leq \rho' < \gamma\lambda_{PQ,1},$$

so $|\rho' - \rho| \leq \max\{\|\boldsymbol{y}\|_2/\tau, \|\boldsymbol{y}'\|_2/\tau\}$ and

$$\|\boldsymbol{h} - \boldsymbol{h}'\|_2 \leq \frac{\|\boldsymbol{y} - \boldsymbol{y}'\|_2}{C_{\gamma,\tau}} + \frac{\max\{\|\boldsymbol{y}\|_2, \|\boldsymbol{y}'\|_2\}\min\{\|\boldsymbol{y}\|_2, \|\boldsymbol{y}'\|_2\}}{\tau C_{\gamma,\tau}^2}$$
$$= \frac{\|\boldsymbol{y} - \boldsymbol{y}'\|_2}{C_{\gamma,\tau}} + \frac{\|\boldsymbol{y}\|_2\|\boldsymbol{y}'\|_2}{\tau C_{\gamma,\tau}^2}.$$

For unconstrained MAVR in optimization (10), we have

$$\|\boldsymbol{h} - \boldsymbol{h}'\|_2 \leq \frac{\|\boldsymbol{y} - \boldsymbol{y}'\|_2}{C_{\gamma,\tau}},$$

since $\rho = \rho' = -1$. $\qquad\square$

A.3. Proof of Theorem 4

Denote by $\boldsymbol{h} = \mathrm{vec}(H)$, $\boldsymbol{y} = \mathrm{vec}(Y)$, $\boldsymbol{h}^* = \mathrm{vec}(H^*)$, $\boldsymbol{e} = \mathrm{vec}(E)$, and $M = \gamma P \otimes Q$. The Kronecker product $P \otimes Q$ is symmetric and positive definite, and then $M^{1/2}$ is a well-defined symmetric and positive-definite matrix. We can know based on $V(H^*) \leq C_h$ that

$$\|M^{1/2}\boldsymbol{h}^*\|_2 = \sqrt{\gamma\boldsymbol{h}^{*\top}(P \otimes Q)\boldsymbol{h}^*} \leq \sqrt{\gamma C_h\|\boldsymbol{h}^*\|_2^2} = \sqrt{\gamma C_h}\|\boldsymbol{h}^*\|_2.$$

Let $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ be two functions extracting the smallest and largest eigenvalues of a matrix. In the following, we will frequently use that $\|A\boldsymbol{v}\|_2 \leq \lambda_{\max}(A)\|\boldsymbol{v}\|_2$ for any symmetric positive-definite matrix A and any vector $\boldsymbol{v}$.

Consider unconstrained MAVR in optimization (10) first. Since $\rho = -1$,

$$\boldsymbol{h} - \boldsymbol{h}^* = (M + I_{nc})^{-1}\boldsymbol{y} - \boldsymbol{h}^*$$
$$= (M + I_{nc})^{-1}(\boldsymbol{h}^* + \boldsymbol{e}) - (M + I_{nc})^{-1}(M + I_{nc})\boldsymbol{h}^*$$
$$= -(M + I_{nc})^{-1}M\boldsymbol{h}^* + (M + I_{nc})^{-1}\boldsymbol{e}.$$

As a consequence,

$$\mathbb{E}\|\boldsymbol{h} - \boldsymbol{h}^*\|_2^2 = \|(M + I_{nc})^{-1}M\boldsymbol{h}^*\|_2^2 + \mathbb{E}\|(M + I_{nc})^{-1}\boldsymbol{e}\|_2^2,$$

since $\mathbb{E}[(M + I_{nc})^{-1}\boldsymbol{e}] = (M + I_{nc})^{-1}\mathbb{E}\boldsymbol{e} = \boldsymbol{0}_{nc}$. Subsequently,

$$\|(M + I_{nc})^{-1}M\boldsymbol{h}^*\|_2 \leq \lambda_{\max}((M + I_{nc})^{-1}M^{1/2}) \cdot \|M^{1/2}\boldsymbol{h}^*\|_2$$
$$\leq \lambda_{\max}((\gamma P \otimes Q + I_{nc})^{-1}(\gamma P \otimes Q)^{1/2}) \cdot \sqrt{\gamma C_h}\|\boldsymbol{h}^*\|_2$$
$$= \sqrt{\gamma C_h}\lambda_{\max}\left(\frac{\sqrt{\gamma}}{\gamma + 1}(\Lambda_{PQ} + I_{nc})^{-1}\Lambda_{PQ}^{1/2}\right)\|\boldsymbol{h}^*\|_2$$
$$\leq \sqrt{C_h}\lambda_{\max}((\Lambda_{PQ} + I_{nc})^{-1}\Lambda_{PQ}^{1/2})\|\boldsymbol{h}^*\|_2$$
$$\leq \frac{1}{2}\sqrt{C_h}\|\boldsymbol{h}^*\|_2,$$

where the last inequality is because the eigenvalues of $(\Lambda_{PQ} + I_{nc})^{-1}\Lambda_{PQ}^{1/2}$ are $\frac{\sqrt{\lambda_{PQ,1}}}{\lambda_{PQ,1}+1}, \ldots, \frac{\sqrt{\lambda_{PQ,nc}}}{\lambda_{PQ,nc}+1}$ and

$$\sup_{\lambda \geq 0} \frac{\sqrt{\lambda}}{\lambda + 1} = \frac{1}{2}.$$

On the other hand,

$$
\begin{aligned}
\mathbb{E}\|(M + I_{nc})^{-1}\boldsymbol{e}\|_2^2 &\leq (\lambda_{\max}((M + I_{nc})^{-1}))^2 \cdot \mathbb{E}\|\boldsymbol{e}\|_2^2 \\
&= \frac{\mathbb{E}[\boldsymbol{e}^\top \boldsymbol{e}]}{(\lambda_{\min}(M + I_{nc}))^2} \\
&\leq \widetilde{l}\sigma_l^2 + \widetilde{u}\sigma_u^2.
\end{aligned}
$$

Hence,

$$\mathbb{E}\|\boldsymbol{h} - \boldsymbol{h}^*\|_2^2 \leq \frac{1}{4}C_h\|\boldsymbol{h}^*\|_2^2 + \widetilde{l}\sigma_l^2 + \widetilde{u}\sigma_u^2,$$

which completes the proof of inequality (20).

Next, consider MAVR in optimization (9). We would have

$$
\begin{aligned}
\boldsymbol{h} - \boldsymbol{h}^* &= (M - \rho I_{nc})^{-1}\boldsymbol{y} - \boldsymbol{h}^* \\
&= (M - \rho I_{nc})^{-1}(\boldsymbol{h}^* + \boldsymbol{e}) - (M - \rho I_{nc})^{-1}(M - \rho I_{nc})\boldsymbol{h}^* \\
&= -(M - \rho I_{nc})^{-1}(M - (\rho + 1)I_{nc})\boldsymbol{h}^* + (M - \rho I_{nc})^{-1}\boldsymbol{e}.
\end{aligned}
$$

In general, $\mathbb{E}[(M - \rho I_{nc})^{-1}\boldsymbol{e}] \neq \boldsymbol{0}_{nc}$ since ρ depends on $\boldsymbol{e}$. Furthermore, $M - (\rho + 1)I_{nc}$ may have negative eigenvalues when $\gamma\lambda_{PQ,1} - 1 < \rho \leq \gamma\lambda_{PQ,1} - C_{\gamma,\tau}$. Taking the expectation of $\|\boldsymbol{h} - \boldsymbol{h}^*\|_2$,

$$
\begin{aligned}
\mathbb{E}\|\boldsymbol{h} - \boldsymbol{h}^*\|_2 &\leq \mathbb{E}\|(M - \rho I_{nc})^{-1}(M - (\rho + 1)I_{nc})\boldsymbol{h}^*\|_2 + \mathbb{E}\|(M - \rho I_{nc})^{-1}\boldsymbol{e}\|_2 \\
&\leq \mathbb{E}\|(M - \rho I_{nc})^{-1}M\boldsymbol{h}^*\|_2 + \mathbb{E}[|\rho + 1|\|(M - \rho I_{nc})^{-1}\boldsymbol{h}^*\|_2] + \mathbb{E}\|(M - \rho I_{nc})^{-1}\boldsymbol{e}\|_2.
\end{aligned}
$$

Subsequently,

$$
\begin{aligned}
\mathbb{E}\|(M - \rho I_{nc})^{-1}M\boldsymbol{h}^*\|_2 &\leq \sup_\rho \lambda_{\max}((M - \rho I_{nc})^{-1}M^{1/2}) \cdot \sqrt{\gamma C_h}\|\boldsymbol{h}^*\|_2 \\
&= \sup_\rho \sqrt{C_h}\lambda_{\max}\left((\Lambda_{PQ} - \rho/\gamma I_{nc})^{-1}\Lambda_{PQ}^{1/2}\right)\|\boldsymbol{h}^*\|_2 \\
&\leq \sqrt{C_h}\|\boldsymbol{h}^*\|_2 \cdot \sup_{\rho \leq \gamma\lambda_{PQ,1} - C_{\gamma,\tau}} \sup_{\lambda \geq \lambda_{PQ,1}} \left(\frac{\sqrt{\lambda}}{\lambda - \rho/\gamma}\right) \\
&\leq \frac{\sqrt{C_h}\gamma\lambda_{PQ,1}}{C_{\gamma,\tau}}\|\boldsymbol{h}^*\|_2.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\mathbb{E}[|\rho + 1|\|(M - \rho I_{nc})^{-1}\boldsymbol{h}^*\|_2] &\leq \mathbb{E}|\rho + 1| \cdot \sup_\rho \lambda_{\max}((M - \rho I_{nc})^{-1})\|\boldsymbol{h}^*\|_2 \\
&\leq \frac{\|\boldsymbol{h}^*\|_2}{C_{\gamma,\tau}} \cdot \mathbb{E}\max\{-\rho - 1, \sup_\rho \rho + 1\} \\
&\leq \frac{\|\boldsymbol{h}^*\|_2}{C_{\gamma,\tau}} \cdot \max\{\mathbb{E}\|\boldsymbol{y}\|_2/\tau - \gamma\lambda_{PQ,1} - 1, \gamma\lambda_{PQ,1} - C_{\gamma,\tau} + 1\} \\
&= \frac{\|\boldsymbol{h}^*\|_2}{C_{\gamma,\tau}} \cdot \max\{\sqrt{\widetilde{l}}/\tau - \gamma\lambda_{PQ,1} - 1, \gamma\lambda_{PQ,1} - C_{\gamma,\tau} + 1\}.
\end{aligned}
$$

where we used the fact that $\sup_\rho \rho$ is independent of $\boldsymbol{e}$, and applied *Jensen's inequality* to obtain that

$$\mathbb{E}\|\boldsymbol{y}\|_2 \leq \sqrt{\mathbb{E}\|\boldsymbol{y}\|_2^2} \leq \sqrt{\widetilde{l}}.$$

In the end,

$$\mathbb{E}\|(M - \rho I_{nc})^{-1}\boldsymbol{e}\|_2 \leq \sup_\rho \lambda_{\max}((M - \rho I_{nc})^{-1}) \cdot \mathbb{E}\|\boldsymbol{e}\|_2$$
$$\leq \frac{\mathbb{E}\sqrt{\boldsymbol{e}^\top \boldsymbol{e}}}{C_{\gamma,\tau}}$$
$$\leq \frac{\sqrt{\mathbb{E}[\boldsymbol{e}^\top \boldsymbol{e}]}}{C_{\gamma,\tau}}$$
$$= \frac{\sqrt{\tilde{l}\sigma_l^2 + \tilde{u}\sigma_u^2}}{C_{\gamma,\tau}},$$

where the third inequality is due to Jensen's inequality. Therefore, inequality (19) follows by combining the three upper bounds of expectations. $\qquad\square$