Supplemental Material on "From Model Selection to Model Averaging: A Comparison for Nested Linear Models"

Wenchao Xu[1] and Xinyu Zhang[2,3]

[1]School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai, China

[2]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

[3]International Institute of Finance, School of Management, University of Science and Technology of China, Hefei, China

We present the supplemental material continued from the main paper, which includes

S.1 Proofs of Theorems 2–6, Lemma 1, and Corollary 1

In the subsequent proofs, all results will be derived on $\mathcal{F}$ when using Assumptions 1–6, Conditions M1–M2, and Conditions A1–A2.

S.1.1 Proof of Theorem 2

We first prove part (i). Let $C > 0$ be a sufficiently large constant. From Assumption 5, there exists a constant $K_C^* = \max\{K_0, \lfloor 2/\bar{\theta}_{\lfloor C\rfloor+1}\rfloor + 1\} > 0$ such that $\theta_{n,\lfloor C\rfloor+1} - 1/n \geq \bar{\theta}_{\lfloor C\rfloor+1}/2 > 0$ for any $n \geq K_C^*$. Since m_n^{**} satisfies $1/n \geq \theta_{n,m_n^{**}+1}$ from (A.2), we have

$$\theta_{n,\lfloor C\rfloor+1} - \theta_{n,m_n^{**}+1} \geq \theta_{n,\lfloor C\rfloor+1} - \frac{1}{n} > 0,$$

which, along with Assumption 3, leads to $m_n^{**} + 1 \geq \lfloor C\rfloor + 2$. This further implies that for any constant $C > 0$, there exists a constant $K_C^* > 0$ such that $m_n^{**} \geq \lfloor C\rfloor + 1 > C$ for any $n \geq K_C^*$, i.e., $\lim_{n\to\infty} m_n^{**} = \infty$. This completes the proof of Theorem 2(i).

Next, we prove part (ii). When $M_n \geq m_n^{**}$, we have $m_n^* = m_n^{**}$, and thus

$$R_n^{\mathrm{MS}}(m_n^*) \geq \mathrm{tr}(\mathbf{P}_{m_n^*}\mathbf{\Omega}) = \mathrm{tr}(\mathbf{P}_{m_n^{**}}\mathbf{\Omega}) \geq c_1\nu_{m_n^{**}} \geq c_1 m_n^{**} \to \infty.$$

When $M_n < m_n^{**}$, we have $m_n^* = M_n$, and thus by (A.2) and Assumptions 2–3,

$$\begin{aligned}
R_n^{\mathrm{MS}}(m_n^*) &= R_n^{\mathrm{MS}}(M_n) = \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{M_n})\boldsymbol{\mu} + \mathrm{tr}(\mathbf{P}_{M_n}\mathbf{\Omega}) \\
&\geq \boldsymbol{\mu}^\top(\mathbf{P}_{m_n^{**}} - \mathbf{P}_{M_n})\boldsymbol{\mu} + \mathrm{tr}(\mathbf{P}_{M_n}\mathbf{\Omega}) \\
&= \sum_{m=M_n+1}^{m_n^{**}} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\} n\theta_{n,m} + \mathrm{tr}(\mathbf{P}_{M_n}\mathbf{\Omega}) \\
&\geq n\theta_{n,m_n^{**}} \sum_{m=M_n+1}^{m_n^{**}} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\} + \mathrm{tr}(\mathbf{P}_{M_n}\mathbf{\Omega}) \\
&\geq \mathrm{tr}\{(\mathbf{P}_{m_n^{**}} - \mathbf{P}_{M_n})\mathbf{\Omega}\} + \mathrm{tr}(\mathbf{P}_{M_n}\mathbf{\Omega})
\end{aligned}$$

$$= \mathrm{tr}(\mathbf{P}_{m_n^{**}}\boldsymbol{\Omega}) \geq c_1\nu_{m_n^{**}} \geq c_1 m_n^{**} \to \infty. \tag{S.1}$$

Therefore, $R_n^{\mathrm{MS}}(m_n^*) \to \infty$ as $n \to \infty$ for any M_n. Combining this fact with Theorem 1, we have $R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \geq R_n^{\mathrm{MS}}(m_n^*)/2 \to \infty$ as $n \to \infty$. This completes the proof of Theorem 2(ii).

S.1.2 Proof of Theorem 3

Under Condition M1, $\lim_{n\to\infty} M_n/m_n^{**} = 0$, which implies that $M_n < m_n^{**}$ when n is large enough, and thus $m_n^* = M_n$. By (A.8), for a sufficiently large n,

$$\begin{aligned}\Delta_n &= \sum_{m=2}^{M_n} \frac{[\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}]^2}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}} \\ &\leq \sum_{m=2}^{M_n} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \\ &= \mathrm{tr}\{(\mathbf{P}_{M_n} - \mathbf{P}_1)\boldsymbol{\Omega}\} \\ &\leq c_2(\nu_{M_n} - \nu_1) \leq c_2 V(M_n - 1), \end{aligned} \tag{S.2}$$

where the last two inequalities are due to Assumptions 2 and 4, respectively. Combining (S.1), (S.2), $\lim_{n\to\infty} M_n/m_n^{**} = 0$, and Theorem 2, we have

$$\limsup_{n\to\infty} \frac{\Delta_n}{R_n^{\mathrm{MS}}(m_n^*)} \leq \frac{c_2 V}{c_1} \lim_{n\to\infty} \frac{M_n - 1}{m_n^{**}} = 0,$$

which yields $\Delta_n = o\{R_n^{\mathrm{MS}}(m_n^*)\}$. This completes the proof of Theorem 3.

S.1.3 Proof of Theorem 4

When Condition M2 holds, we consider two scenarios to prove this theorem: $M_n \geq m_n^{**}$ and $\underline{c} \leq M_n/m_n^{**} < 1$, for any sufficiently large n.

(i) $M_n \geq m_n^{**}$ for any sufficiently large n. In this case, $m_n^* = m_n^{**}$ satisfies (A.4). When Condition A1 holds, we first examine the order of the optimal risk of MS. Let $s_n^* = \max\{s : \lfloor k^s(m_n^*+1)\rfloor \leq d_n, s = 0, 1, \ldots\}$. The first term in (A.3) is upper bounded by

$$\begin{aligned}&\boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{m_n^*})\boldsymbol{\mu} \\ &= \sum_{m=m_n^*+1}^{q_n} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} = \sum_{m=m_n^*+1}^{d_n} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\end{aligned}$$

$$
\begin{aligned}
&= \sum_{s=0}^{s_n^*-1} \sum_{m=\lfloor k^s(m_n^*+1)\rfloor}^{\lfloor k^{s+1}(m_n^*+1)\rfloor-1} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \sum_{m=\lfloor k^{s_n^*}(m_n^*+1)\rfloor}^{d_n} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} \\
&= \sum_{s=0}^{s_n^*-1} \sum_{m=\lfloor k^s(m_n^*+1)\rfloor}^{\lfloor k^{s+1}(m_n^*+1)\rfloor-1} n\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\theta_{n,m} + \sum_{m=\lfloor k^{s_n^*}(m_n^*+1)\rfloor}^{d_n} n\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\theta_{n,m} \\
&\le \sum_{s=0}^{s_n^*-1} \theta_{n,\lfloor k^s(m_n^*+1)\rfloor} \sum_{m=\lfloor k^s(m_n^*+1)\rfloor}^{\lfloor k^{s+1}(m_n^*+1)\rfloor-1} n\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \\
&\quad + \theta_{n,\lfloor k^{s_n^*}(m_n^*+1)\rfloor} \sum_{m=\lfloor k^{s_n^*}(m_n^*+1)\rfloor}^{d_n} n\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \\
&\le n\theta_{n,m_n^*+1} \sum_{s=0}^{s_n^*-1} \eta^s \mathrm{tr}\{(\mathbf{P}_{\lfloor k^{s+1}(m_n^*+1)\rfloor-1} - \mathbf{P}_{\lfloor k^s(m_n^*+1)\rfloor-1})\boldsymbol{\Omega}\} \\
&\quad + n\theta_{n,m_n^*+1}\eta^{s_n^*}\mathrm{tr}\{(\mathbf{P}_{d_n} - \mathbf{P}_{\lfloor k^{s_n^*}(m_n^*+1)\rfloor-1})\boldsymbol{\Omega}\} \\
&\le c_2 \sum_{s=0}^{s_n^*-1} \eta^s(\nu_{\lfloor k^{s+1}(m_n^*+1)\rfloor} - \nu_{\lfloor k^s(m_n^*+1)\rfloor}) + c_2\eta^{s_n^*}(\nu_{d_n} - \nu_{\lfloor k^{s_n^*}(m_n^*+1)\rfloor-1}) \\
&\le c_2 V \sum_{s=0}^{s_n^*} \eta^s(\lfloor k^{s+1}(m_n^*+1)\rfloor - \lfloor k^s(m_n^*+1)\rfloor) \\
&\sim c_2V(k-1)(m_n^*+1)\sum_{s=0}^{s_n^*}(k\eta)^s \asymp m_n^* \asymp \mathrm{tr}(\mathbf{P}_{m_n^*}\boldsymbol{\Omega}).
\end{aligned}
$$

In this progression, the first equality follows from the fact that $\boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{q_n})\boldsymbol{\mu} = 0$; the first inequality follows from Assumption 3; the second inequality follows from $\theta_{n,\lfloor k^s(m_n^*+1)\rfloor}/\theta_{n,m_n^*+1} \le \eta^s$ for a sufficiently large n, which can be obtained by Condition A1 and Theorem 2; and the last two inequalities follow from (A.4) and Assumption 4 respectively. Thus, the order of the optimal risk of MS satisfies $R_n^{\mathrm{MS}}(m_n^*) \asymp \mathrm{tr}(\mathbf{P}_{m_n^*}\boldsymbol{\Omega})$.

Next, we prove that the potential advantage Δ_n of MA over MS has the same order as $R_n^{\mathrm{MS}}(m_n^*)$ under Condition A1. Define $t_n = \min\{t \in \mathbb{N}\colon \lfloor kt \rfloor \ge m_n^*+1\}$. Then it follows from Theorem 2 and Peng and Yang (2022) that $\lim_{n\to\infty} t_n = \infty$, $\lfloor kt_n \rfloor \sim m_n^*$, and $t_n \sim m_n^*/k$. The first term in (A.8) can be lower bounded by

$$
\begin{aligned}
&\sum_{m=2}^{m_n^*}\left[\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} - \frac{\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}}\right] \\
&\quad \ge \mathrm{tr}\{(\mathbf{P}_{m_n^*} - \mathbf{P}_1)\boldsymbol{\Omega}\} - \sum_{m=2}^{\lfloor kt_n \rfloor} \frac{\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}}
\end{aligned}
$$

$$
\begin{aligned}
&\geq \operatorname{tr}\{(\mathbf{P}_{m_n^*} - \mathbf{P}_1)\boldsymbol{\Omega}\} - \sum_{m=2}^{t_n} \operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} - \sum_{m=t_n+1}^{\lfloor kt_n \rfloor} \frac{\operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}}{1 + 1/(n\theta_{n,m})} \\
&\geq \operatorname{tr}\{(\mathbf{P}_{m_n^*} - \mathbf{P}_{t_n})\boldsymbol{\Omega}\} - \frac{1}{1 + 1/(n\theta_{n,t_n})} \operatorname{tr}\{(\mathbf{P}_{\lfloor kt_n \rfloor} - \mathbf{P}_{t_n})\boldsymbol{\Omega}\}, \\
&\geq \operatorname{tr}\{(\mathbf{P}_{m_n^*} - \mathbf{P}_{t_n})\boldsymbol{\Omega}\} - \frac{1}{1+\delta} \operatorname{tr}\{(\mathbf{P}_{\lfloor kt_n \rfloor} - \mathbf{P}_{t_n})\boldsymbol{\Omega}\} \\
&= \frac{1}{1+\delta} \operatorname{tr}\big[\{(1+\delta)\mathbf{P}_{m_n^*} - \mathbf{P}_{\lfloor kt_n \rfloor} - \delta \mathbf{P}_{t_n}\}\boldsymbol{\Omega}\big] \qquad \text{(S.3)}
\end{aligned}
$$

where the third inequality follows from Assumption 3, and the last inequality follows from the following fact

$$
\frac{1}{1 + 1/(n\theta_{n,t_n})} \leq \frac{1}{1 + \delta/(n\theta_{n,\lfloor kt_n \rfloor})} \leq \frac{1}{1 + \delta/(n\theta_{n,m_n^*+1})} \leq \frac{1}{1+\delta},
$$

which can be derived by (A.4) and Condition A1. Since $\nu_{m_n^*} \sim \nu_{\lfloor kt_n \rfloor}$, it is easy to show that $(1+\delta)\mathbf{P}_{m_n^*} - \mathbf{P}_{\lfloor kt_n \rfloor} - \delta \mathbf{P}_{t_n}$ is positive semi-definite for sufficiently large n. By Assumption 2 and the fact that $\operatorname{tr}(\mathbf{AB}) \geq \lambda_{\min}(\mathbf{A})\operatorname{tr}(\mathbf{B})$ for symmetric matrix $\mathbf{A}$ and positive semi-definite matrix $\mathbf{B}$ (Bernstein, 2005, Proposition 8.4.13), we have

$$
\begin{aligned}
&\frac{1}{1+\delta} \operatorname{tr}\big[\{(1+\delta)\mathbf{P}_{m_n^*} - \mathbf{P}_{\lfloor kt_n \rfloor} - \delta \mathbf{P}_{t_n}\}\boldsymbol{\Omega}\big] \\
&\quad \geq \frac{c_1}{1+\delta}\{(1+\delta)\nu_{m_n^*} - \nu_{\lfloor kt_n \rfloor} - \delta \nu_{t_n}\} \\
&\quad \geq \frac{c_1}{1+\delta}(\nu_{m_n^*} - \nu_{\lfloor kt_n \rfloor}) + \frac{c_1 \delta}{1+\delta}(m_n^* - t_n) \\
&\quad \sim \frac{(k-1)c_1\delta}{k(1+\delta)} m_n^* \asymp \operatorname{tr}(\mathbf{P}_{m_n^*}\boldsymbol{\Omega}), \qquad \text{(S.4)}
\end{aligned}
$$

where the last line is due to $\nu_{m_n^*} \sim \nu_{\lfloor kt_n \rfloor}$ and $t_n \sim m_n^*/k$. From (A.8), we see

$$
R_n^{\mathrm{MS}}(m_n^*) \geq \Delta_n \geq \sum_{m=2}^{m_n^*} \left[\operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} - \frac{\operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}} \right],
$$

which, along with (S.3) and (S.4), implies $\Delta_n \asymp R_n^{\mathrm{MS}}(m_n^*)$. This completes the proof of the result under Condition A1.

When Condition A2 holds, we examine $\Delta_n = o\{R_n^{\mathrm{MS}}(m_n^*)\}$. Let $2/m_n^* < k' < 1$. The first term in (A.8) is upper bounded by

$$
\sum_{m=2}^{m_n^*} \left[\operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} - \frac{\operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \operatorname{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}} \right]
$$

$$
\begin{aligned}
&\le \operatorname{tr}\{(\mathbf{P}_{m_n^*}-\mathbf{P}_1)\boldsymbol{\Omega}\} - \sum_{m=2}^{\lfloor k' m_n^* \rfloor} \frac{\operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\}}{1+1/(n\theta_{n,m})} \\
&\le \operatorname{tr}\{(\mathbf{P}_{m_n^*}-\mathbf{P}_1)\boldsymbol{\Omega}\} - \frac{1}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})} \sum_{m=2}^{\lfloor k' m_n^* \rfloor} \operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\} \\
&= \operatorname{tr}\{(\mathbf{P}_{m_n^*}-\mathbf{P}_1)\boldsymbol{\Omega}\} - \frac{1}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})} \operatorname{tr}\{(\mathbf{P}_{\lfloor k' m_n^* \rfloor}-\mathbf{P}_1)\boldsymbol{\Omega}\} \\
&= \operatorname{tr}\left\{\left(\mathbf{P}_{m_n^*}-\mathbf{P}_1-\frac{\mathbf{P}_{\lfloor k' m_n^* \rfloor}-\mathbf{P}_1}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})}\right)\boldsymbol{\Omega}\right\}. \qquad \text{(S.5)}
\end{aligned}
$$

Observe that

$$
\mathbf{P}_{m_n^*}-\mathbf{P}_1-\frac{\mathbf{P}_{\lfloor k' m_n^* \rfloor}-\mathbf{P}_1}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})} = \frac{\mathbf{P}_{m_n^*}-\mathbf{P}_{\lfloor k' m_n^* \rfloor}+(\mathbf{P}_{m_n^*}-\mathbf{P}_1)/(n\theta_{n,\lfloor k' m_n^* \rfloor})}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})}
$$

is a positive semi-definite matrix. By the fact that $\operatorname{tr}(\mathbf{AB}) \le \lambda_{\max}(\mathbf{A})\operatorname{tr}(\mathbf{B})$ for symmetric matrix $\mathbf{A}$ and positive semi-definite matrix $\mathbf{B}$ (Bernstein, 2005, Proposition 8.4.13), we have

$$
\begin{aligned}
&\operatorname{tr}\left\{\left(\mathbf{P}_{m_n^*}-\mathbf{P}_1-\frac{\mathbf{P}_{\lfloor k' m_n^* \rfloor}-\mathbf{P}_1}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})}\right)\boldsymbol{\Omega}\right\} \\
&\le \frac{c_2}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})}\operatorname{tr}\left(\mathbf{P}_{m_n^*}-\mathbf{P}_{\lfloor k' m_n^* \rfloor}+\frac{\mathbf{P}_{m_n^*}-\mathbf{P}_1}{n\theta_{n,\lfloor k' m_n^* \rfloor}}\right) \\
&= \frac{c_2}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})}\left(\nu_{m_n^*}-\nu_{\lfloor k' m_n^* \rfloor}+\frac{\nu_{m_n^*}-\nu_1}{n\theta_{n,\lfloor k' m_n^* \rfloor}}\right) \\
&\le \frac{c_2 V}{1+1/(n\theta_{n,\lfloor k' m_n^* \rfloor})}\left(m_n^*-\lfloor k' m_n^* \rfloor+\frac{m_n^*-1}{n\theta_{n,\lfloor k' m_n^* \rfloor}}\right) \\
&\le c_2 V\left\{m_n^*-\lfloor k' m_n^* \rfloor+\frac{\theta_{n,m_n^*}}{\theta_{n,\lfloor k' m_n^* \rfloor}}(m_n^*-1)\right\}, \qquad \text{(S.6)}
\end{aligned}
$$

where the second inequality follows from Assumption 4. Since $\lim_{n\to\infty}\theta_{n,m_n^*}/\theta_{n,\lfloor k' m_n^* \rfloor}=0$ for any $k'<1$ under Condition A2 and Theorem 2, we have

$$
\left\{m_n^*-\lfloor k' m_n^* \rfloor+\frac{\theta_{n,m_n^*}}{\theta_{n,\lfloor k' m_n^* \rfloor}}(m_n^*-1)\right\}\Big/ m_n^* = 1-\frac{\lfloor k' m_n^* \rfloor}{m_n^*}+\frac{\theta_{n,m_n^*}}{\theta_{n,\lfloor k' m_n^* \rfloor}}\left(1-\frac{1}{m_n^*}\right)\to 1-k',
$$

which along with (S.5) and (S.6), yields that

$$
\sum_{m=2}^{m_n^*}\left[\operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\}-\frac{\operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\}\boldsymbol{\mu}^\top(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\mu}}{\boldsymbol{\mu}^\top(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\mu}+\operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\}}\right]=O\{(1-k')m_n^*\}.
$$

Due to the arbitrariness of k' and the fact $\operatorname{tr}(\mathbf{P}_{m_n^*}\boldsymbol{\Omega}) \asymp m_n^*$, letting $k'\to 1$, we can obtain the first term of (A.8):

$$
\sum_{m=2}^{m_n^*}\left[\operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\}-\frac{\operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\}\boldsymbol{\mu}^\top(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\mu}}{\boldsymbol{\mu}^\top(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\mu}+\operatorname{tr}\{(\mathbf{P}_m-\mathbf{P}_{m-1})\boldsymbol{\Omega}\}}\right]=o\{\operatorname{tr}(\mathbf{P}_{m_n^*}\boldsymbol{\Omega})\}. \qquad \text{(S.7)}
$$

Next, we consider the order of the second term of (A.8). Choose $k > 1$. We have

$$\sum_{m=m_n^*+1}^{M_n} \frac{\{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\}^2}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}}$$
$$= \sum_{m=m_n^*+1}^{\lfloor k(m_n^*+1)\rfloor} \frac{n\theta_{n,m}}{1+1/(n\theta_{n,m})}\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} + \sum_{m=\lfloor k(m_n^*+1)\rfloor+1}^{\min\{M_n,d_n\}} \frac{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{1+1/(n\theta_{n,m})}. \tag{S.8}$$

The first term of (S.8) is upper bounded by

$$\begin{aligned}
&\sum_{m=m_n^*+1}^{\lfloor k(m_n^*+1)\rfloor} \frac{n\theta_{n,m}}{1+1/(n\theta_{n,m})}\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \\
&\le \frac{n\theta_{n,m_n*+1}}{1+1/(n\theta_{n,m_n^*+1})} \sum_{m=m_n^*+1}^{\lfloor k(m_n^*+1)\rfloor} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \\
&\le \frac{1}{2}\mathrm{tr}\{(\mathbf{P}_{\lfloor k(m_n^*+1)\rfloor} - \mathbf{P}_{m_n^*})\boldsymbol{\Omega}\} \\
&\le \frac{c_2}{2}(\nu_{\lfloor k(m_n^*+1)\rfloor} - \nu_{m_n^*}) \\
&\le \frac{c_2}{2}V(\lfloor k(m_n^*+1)\rfloor - m_n^*),
\end{aligned}$$

where the first two inequalities follow from Assumption 3 and (A.4), respectively, and the last inequality follows from Assumption 4. Using Theorem 2, as $n \to \infty$,

$$\frac{\lfloor k(m_n^*+1)\rfloor - m_n^*}{m_n^*} = \frac{\lfloor k(m_n^*+1)\rfloor}{m_n^*} - 1 \to k-1.$$

Therefore,

$$\sum_{m=m_n^*+1}^{\lfloor k(m_n^*+1)\rfloor} \frac{n\theta_{n,m}}{1+1/(n\theta_{n,m})}\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} = O\{(k-1)m_n^*\} = O\{(k-1)\mathrm{tr}(\mathbf{P}_{m_n^*}\boldsymbol{\Omega})\}. \tag{S.9}$$

The second term of (S.8) can be upper bounded by

$$\begin{aligned}
&\sum_{m=\lfloor k(m_n^*+1)\rfloor+1}^{\min\{M_n,d_n\}} \frac{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{1+1/(n\theta_{n,m})} \\
&\le \frac{1}{1+1/(n\theta_{n,\lfloor k(m_n^*+1)\rfloor})} \sum_{m=\lfloor k(m_n^*+1)\rfloor+1}^{\min\{M_n,d_n\}} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} \\
&\le \frac{1}{1+\theta_{n,m_n^*+1}/\theta_{n,\lfloor k(m_n^*+1)\rfloor}}\boldsymbol{\mu}^\top(\mathbf{P}_{\min\{M_n,d_n\}} - \mathbf{P}_{\lfloor k(m_n^*+1)\rfloor})\boldsymbol{\mu} \\
&\le \frac{1}{1+\theta_{n,m_n^*+1}/\theta_{n,\lfloor k(m_n^*+1)\rfloor}}\boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{m_n^*})\boldsymbol{\mu}
\end{aligned}$$

$$= o\{\boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{m_n^*})\boldsymbol{\mu}\}, \tag{S.10}$$

where the first two inequalities follow from Assumption 3 and (A.4), respectively, and the last inequality follows from the following fact:

$$\begin{aligned}\boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{m_n^*})\boldsymbol{\mu} &= \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{\min\{M_n,d_n\}})\boldsymbol{\mu} + \boldsymbol{\mu}^\top(\mathbf{P}_{\min\{M_n,d_n\}} - \mathbf{P}_{\lfloor k(m_n^*+1)\rfloor})\boldsymbol{\mu} \\ &\quad + \boldsymbol{\mu}^\top(\mathbf{P}_{\lfloor k(m_n^*+1)\rfloor} - \mathbf{P}_{m_n^*})\boldsymbol{\mu} \\ &\geq \boldsymbol{\mu}^\top(\mathbf{P}_{\min\{M_n,d_n\}} - \mathbf{P}_{\lfloor k(m_n^*+1)\rfloor})\boldsymbol{\mu}.\end{aligned}$$

The last equality of (S.10) follows from the fact that $\lim_{n\to\infty}\theta_{n,\lfloor k(m_n^*+1)\rfloor}/\theta_{n,m_n^*+1} = 0$ for any $k > 1$ under Condition A2. Combining (S.8), (S.9), and (S.10), we have

$$\sum_{m=m_n^*+1}^{M_n} \frac{\{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\}^2}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}} = O\{(k-1)\mathrm{tr}(\mathbf{P}_{m_n^*}\boldsymbol{\Omega})\} + o\{\boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{m_n^*})\boldsymbol{\mu}\}.$$

Duo to the arbitrariness of k, letting $k \to 1$, we have

$$\sum_{m=m_n^*+1}^{M_n} \frac{\{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\}^2}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}} = o\big\{R_n^{\mathrm{MS}}(m_n^*)\big\},$$

which, along with (A.8) and (S.7), leads to $\Delta_n = o\{R_n^{\mathrm{MS}}(m_n^*)\}$. This completes the proof of the result under Condition A2.

(ii) $\underline{c} \leq M_n/m_n^{**} < 1$ for any sufficiently large n. In this case, $m_n^* = M_n \asymp m_n^{**}$. When Condition A1 holds, there exists a finite positive integer τ_1 such that $k^{-\tau_1} \leq \underline{c}$. Therefore,

$$\theta_{n,m_n^*+1} = \theta_{n,M_n+1} \leq \theta_{n,\lfloor \underline{c} m_n^{**}\rfloor+1} \leq \theta_{n,\lfloor k^{-\tau_1}(m_n^{**}+1)\rfloor} \leq \delta^{-\tau_1}\theta_{n,m_n^{**}+1} \leq \delta^{-\tau_1}\frac{1}{n}, \tag{S.11}$$

where the last inequality is due to (A.2). Then, using the same arguments in (i) and (S.11), it is easy to prove the result under Condition A1. When Condition A2 holds, we can also obtain (S.7), which along with the fact that the second term of (A.8) equals 0, yields the result under Condition A2. This completes the proof of Theorem 4 under (ii).

S.1.4 Proof of Lemma 1

From Assumption 7, we know that for any small $0 < \epsilon < 1$, there exists a constant $K_\epsilon > 0$ which does not depend on m, such that $0 < 1 - \epsilon \leq \theta_{n,m}/\theta_m^* \leq 1 + \epsilon$ holds uniformly in $m = 1, \ldots, d_n$ and $n \geq K_\epsilon$.

(i) When Condition B1 holds, there exist constants $k > 1$ and $0 < \delta^* \leq \eta^* < 1$ with $k\eta^* < 1$ such that for a sufficiently large n,

$$\frac{1-\epsilon}{1+\epsilon}\delta^* \leq \frac{\theta_{n,\lfloor kl_n \rfloor}}{\theta_{n,l_n}} = \frac{\theta_{n,\lfloor kl_n \rfloor}}{\theta^*_{\lfloor kl_n \rfloor}} \times \frac{\theta^*_{\lfloor kl_n \rfloor}}{\theta^*_{l_n}} \times \frac{\theta^*_{l_n}}{\theta_{n,l_n}} \leq \frac{1+\epsilon}{1-\epsilon}\eta^*.$$

Let $\delta = \frac{1-\epsilon}{1+\epsilon}\delta^*$ and $\eta = \frac{1+\epsilon}{1-\epsilon}\eta^*$. Since $\lim_{\epsilon\to 0}\frac{1+\epsilon}{1-\epsilon} = 1$, we can choose a small enough $\epsilon > 0$ such that $0 < \delta \leq \eta < 1$ and $k\eta < 1$. Therefore, Condition B1 implies Condition A1.

(ii) When Condition B2 holds, for every constant $k > 1$ and every integer sequence $\{l_n\}$ satisfied $\lim_{n\to\infty} l_n = \infty$,

$$\lim_{n\to\infty}\frac{\theta_{n,\lfloor kl_n \rfloor}}{\theta_{n,l_n}} = \lim_{n\to\infty}\left\{\frac{\theta_{n,\lfloor kl_n \rfloor}}{\theta^*_{\lfloor kl_n \rfloor}} \times \frac{\theta^*_{\lfloor kl_n \rfloor}}{\theta^*_{l_n}} \times \frac{\theta^*_{l_n}}{\theta_{n,l_n}}\right\} \leq \frac{1+\epsilon}{1-\epsilon}\lim_{n\to\infty}\frac{\theta^*_{\lfloor kl_n \rfloor}}{\theta^*_{l_n}} = 0.$$

Therefore, Condition B2 implies Condition A2.

S.1.5 Proof of Corollary 1

From Theorem 1, $1/2 \leq R_n^{\mathrm{MA}}(\mathbf{w}_n^*)/R_n^{\mathrm{MS}}(m_n^*) \leq 1$ for any sufficiently large n. Since $R_n^{\mathrm{MS}}(\widehat{m})/R_n^{\mathrm{MS}}(m_n^*) = 1 + o_p(1)$ and $R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})/R_n^{\mathrm{MA}}(\mathbf{w}_n^*) = 1 + o_p(1)$, we have when n is large enough,

$$\frac{1}{2}\{1 + o_p(1)\} \leq \frac{R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MS}}(\widehat{m})} = \frac{R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}\frac{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}{R_n^{\mathrm{MS}}(m_n^*)}\frac{R_n^{\mathrm{MS}}(m_n^*)}{R_n^{\mathrm{MS}}(\widehat{m})} \leq 1 + o_p(1),$$

which yields that $R_n^{\mathrm{MA}}(\widehat{\mathbf{w}}) \asymp_p R_n^{\mathrm{MS}}(\widehat{m})$. Observe that

$$\frac{R_n^{\mathrm{MS}}(\widehat{m}) - R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MS}}(\widehat{m})} = 1 - \frac{R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}\frac{R_n^{\mathrm{MS}}(m_n^*)}{R_n^{\mathrm{MS}}(\widehat{m})} + \frac{R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}\frac{\Delta_n}{R_n^{\mathrm{MS}}(m_n^*)}\frac{R_n^{\mathrm{MS}}(m_n^*)}{R_n^{\mathrm{MS}}(\widehat{m})}. \quad \text{(S.12)}$$

Under Conditions M2 and A1, from Theorem 4, $\Delta_n/R_n^{\mathrm{MS}}(m_n^*) \geq c^*$ for some $c^* \in (0, 1/2]$ and any sufficiently large n. Therefore, when n is large enough,

$$\begin{aligned} 1 \geq \left|\frac{R_n^{\mathrm{MS}}(\widehat{m}) - R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MS}}(\widehat{m})}\right| &\geq \frac{R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}\frac{\Delta_n}{R_n^{\mathrm{MS}}(m_n^*)}\frac{R_n^{\mathrm{MS}}(m_n^*)}{R_n^{\mathrm{MS}}(\widehat{m})} - \left|1 - \frac{R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}\frac{R_n^{\mathrm{MS}}(m_n^*)}{R_n^{\mathrm{MS}}(\widehat{m})}\right| \\ &\geq c^*\{1 + o_p(1)\} - o_p(1) = c^*\{1 + o_p(1)\}, \end{aligned}$$

which leads to $R_n^{\mathrm{MS}}(\widehat{m}) - R_n^{\mathrm{MA}}(\widehat{\mathbf{w}}) \asymp_p R_n^{\mathrm{MS}}(\widehat{m})$. Under Condition M1 or Conditions M2 and A2, $\lim_{n\to\infty}\Delta_n/R_n^{\mathrm{MS}}(m_n^*) = 0$ from Theorems 3 and 4. Therefore, by (S.12), we have

$$\frac{R_n^{\mathrm{MS}}(\widehat{m}) - R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})}{R_n^{\mathrm{MS}}(\widehat{m})} \xrightarrow{p} 0,$$

which implies that $R_n^{\mathrm{MS}}(\widehat{m}) - R_n^{\mathrm{MA}}(\widehat{\mathbf{w}}) = o_p\{R_n^{\mathrm{MS}}(\widehat{m})\}$ or $R_n^{\mathrm{MS}}(\widehat{m}) \sim_p R_n^{\mathrm{MA}}(\widehat{\mathbf{w}})$. This completes the proof of Corollary 1.

S.1.6 Proof of Theorem 5

From (A.5) and Assumption 3, it is easy to see that the risk of the optimal MA estimator without the total weight constraint is

$$R_n^{\mathrm{MA}}(\widetilde{\mathbf{w}}_n^*) = \sum_{m=1}^{M_n} \frac{\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}} + \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{M_n})\boldsymbol{\mu},$$

which along with (A.7) and Assumption 2, yields that

$$R_n^{\mathrm{MA}}(\mathbf{w}_n^*) - R_n^{\mathrm{MA}}(\widetilde{\mathbf{w}}_n^*) = \frac{\{\mathrm{tr}(\mathbf{P}_1\boldsymbol{\Omega})\}^2}{\boldsymbol{\mu}^\top\mathbf{P}_1\boldsymbol{\mu} + \mathrm{tr}(\mathbf{P}_1\boldsymbol{\Omega})} \leq \mathrm{tr}(\mathbf{P}_1\boldsymbol{\Omega}) < c_2\nu_1.$$

Furthermore, if Assumptions 4–6 hold, we have $R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \to \infty$ from Theorem 2(ii). Therefore, $R_n^{\mathrm{MA}}(\mathbf{w}_n^*) - R_n^{\mathrm{MA}}(\widetilde{\mathbf{w}}_n^*) = o\{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)\}$, which completes the proof of Theorem 5.

S.1.7 Two Lemmas and Their Proofs

Before giving the proof of Theorems 6, we prove two lemmas. Let $\lceil a \rceil$ denote the least integer greater than or equal to $a \in \mathbb{R}$. We first present the following lemma on an expression of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)$.

Lemma S.1. *Suppose that Assumptions 3 and 6 hold. For any sufficiently large n, the optimal risk of MA restricted to $\mathcal{W}_n(N)$ is given by*

$$\begin{aligned} R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) &= \mathrm{tr}(\mathbf{P}_1\boldsymbol{\Omega}) + \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{M_n})\boldsymbol{\mu} \\ &\quad + \sum_{i=i_{n,N}+1}^{N} \sum_{m=m_n(\frac{2i+1}{2N})+1}^{m_n(\frac{2i-1}{2N})} \left[\left(\frac{i}{N}\right)^2 \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} + \left(1 - \frac{i}{N}\right)^2 \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\right] \\ &\quad + \sum_{m=m_n(\frac{2i_{n,N}+1}{2N})+1}^{M_n} \left[\left(\frac{i_{n,N}}{N}\right)^2 \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} + \left(1 - \frac{i_{n,N}}{N}\right)^2 \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\right], \end{aligned}$$

where $i_{n,N} = \lceil N\gamma_{n,M_n}^* - \frac{1}{2} \rceil$, and $m_n(z)$ for $z \in (\gamma_{n,q_n}^*, 1)$ is an integer in $\{1, \ldots, q_n\}$ satisfying

$$\theta_{n,m_n(z)} > \frac{z}{(1-z)n} \geq \theta_{n,m_n(z)+1}, \tag{S.13}$$

and $m_n(z_0) = 1$ for any $z_0 \geq 1$.

Proof. Since $\mathbf{w} \in \mathcal{W}_n(N)$, we have $\gamma_m = \sum_{j=m}^{M_n} w_j \in \{0, 1/N, 2/N, \ldots, 1\}$. Observe that

$$\begin{aligned} f_m(\gamma_m) &\equiv \gamma_m^2 \Big[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\Big] - 2\gamma_m \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} \\ &= \Big[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\Big](\gamma_m - \gamma_{n,m}^*)^2 + \gamma_{n,m}^* \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}, \end{aligned}$$

where $\gamma_{n,m}^*$ is defined in (A.6). Since $\{\gamma_{n,m}^*\}_{m=1}^{M_n}$ is nonincreasing, it is easy to see that

$$\min_{\gamma_m \in \{0,1/N,2/N,\ldots,1\}} f_m(\gamma_m) = f_m\left(\frac{i}{N}\right), \quad \text{when } \frac{2i-1}{2N} < \gamma_{n,m}^* \leq \frac{2i+1}{2N}, \quad i = 0, \ldots, N.$$

Therefore, from (A.5), we have

$$\begin{aligned} R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) &= \mathrm{tr}(\mathbf{P}_1\boldsymbol{\Omega}) + \sum_{m=2}^{M_n} \min_{\gamma_m \in \{0,1/N,2/N,\ldots,1\}} f_m(\gamma_m) + \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{M_n})\boldsymbol{\mu} \\ &= \mathrm{tr}(\mathbf{P}_1\boldsymbol{\Omega}) + \sum_{m=2}^{M_n}\sum_{i=0}^{N} f_m\left(\frac{i}{N}\right) \mathbf{1}\left\{\frac{2i-1}{2N} < \gamma_{n,m}^* \leq \frac{2i+1}{2N}\right\} + \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{M_n})\boldsymbol{\mu} \\ &= \mathrm{tr}(\mathbf{P}_1\boldsymbol{\Omega}) + \sum_{m=2}^{M_n}\sum_{i=0}^{N} \left[\left(\frac{i}{N}\right)^2 \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} + \left(1 - \frac{i}{N}\right)^2 \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\right] \\ &\quad \times \mathbf{1}\left\{\frac{2i-1}{2N} < \gamma_{n,m}^* \leq \frac{2i+1}{2N}\right\} + \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{M_n})\boldsymbol{\mu}, \end{aligned} \tag{S.14}$$

where $\mathbf{1}\{\cdot\}$ denotes the usual indicator function. By the definition of $m_n(z)$ in (S.13), we have $\frac{2i-1}{2N} < \gamma_{n,m}^* \leq \frac{2i+1}{2N}$ if and only if $m_n(\frac{2i+1}{2N}) + 1 \leq m \leq m_n(\frac{2i-1}{2N})$ for $i = i_{n,N}+1, \ldots, N$ and $\frac{2i_{n,N}-1}{2N} < \gamma_{n,m}^* \leq \frac{2i_{n,N}+1}{2N}$ if and only if $m_n(\frac{2i_{n,N}+1}{2N}) + 1 \leq m \leq M_n$, where

$$i_{n,N} = \min\left\{i = 0, 1, \ldots, N \colon \gamma_{n,M_n}^* \leq \frac{2i+1}{2N}\right\} = \left\lceil N\gamma_{n,M_n}^* - \frac{1}{2}\right\rceil.$$

Combining the above fact with (S.14), it is easy to obtain the expression of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)$ in Lemma S.1. Moreover, we can obtain another expression of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)$ as follows:

$$\begin{aligned} R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) = R_n^{\mathrm{MA}}(\mathbf{w}_n^*) + \sum_{m=2}^{M_n}\Bigg(&\Big[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\Big] \\ &\times \sum_{i=0}^{N}\left(\frac{i}{N} - \gamma_{n,m}^*\right)^2 \mathbf{1}\left\{\frac{2i-1}{2N} < \gamma_{n,m}^* \leq \frac{2i+1}{2N}\right\}\Bigg). \end{aligned} \tag{S.15}$$

This completes the proof of Lemma S.1. □

Note that $m_n(1/2) = m_n^{**}$. Next, we present some elementary properties of $m_n(z)$ in the following lemma.

Lemma S.2. *Suppose that Assumptions 3 and 5 hold. Then,* $m_n(z)$ *for* $z \in (\gamma^*_{n,q_n}, 1)$ *defined in Lemma S.1 satisfies the following properties.*

(i) $m_n(z)$ *is a nonincreasing function in* z*;* $\lim_{n\to\infty} m_n(z) = \infty$ *for any fixed* $z \in (\gamma^*_{n,q_n}, 1)$*.*

(ii) If there exist constants $k > 1$*,* $\eta < 1$*, and* $K > 1$ *such that* $\theta_{n,\lfloor k l_n \rfloor}/\theta_{n,l_n} \leq \eta$ *for any* $n \geq K$ *and any integer sequence* $\{l_n\}$ *satisfying* $\lim_{n\to\infty} l_n = \infty$*, then* $m_n(z_1) \asymp m_n(z_2)$ *for any* $\gamma^*_{n,q_n} < z_1 \neq z_2 < 1$*.*

Proof. The results of (i) are easily shown by Assumption 1 and arguments similar to those in the proof of Lemma 2. Next, we shall prove (ii). Without loss of generality, we assume $z_1 < z_2$, from which it follows that $m_n(z_1) \geq m_n(z_2)$. Observe there exists an integer $s > 0$ such that $\frac{z_1}{1-z_1} \geq \frac{z_2}{1-z_2}\eta^s$. Then, by the definition of $m_n(\xi)$, we have

$$\theta_{n,m_n(z_1)} > \frac{z_1}{(1-z_1)n} \geq \frac{z_2}{(1-z_2)n}\eta^s \geq \eta^s \theta_{n,m_n(z_2)+1} \geq \theta_{n,\lfloor k^s(m_n(z_2)+1)\rfloor}. \tag{S.16}$$

Thus, $m_n(z_1) < \lfloor k^s(m_n(z_2)+1)\rfloor$, which, along with $m_n(z_1) \geq m_n(z_2)$, yields that $m_n(z_1) \asymp m_n(z_2)$. This completes the proof of Lemma S.2. □

S.1.8 Proof of Theorem 6

Observe that

$$\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} = \frac{\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}}{\gamma^*_{n,m}} = \frac{\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}}{1-\gamma^*_{n,m}}, \tag{S.17}$$

which, along with (S.15), yields that

$$\begin{aligned}
&R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,N}) - R_n^{\mathrm{MA}}(\mathbf{w}^*_n) \\
&= \sum_{m=2}^{m_n^*} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \sum_{i=0}^{N} \frac{\left(\frac{i}{N} - \gamma^*_{n,m}\right)^2}{1-\gamma^*_{n,m}} \mathbf{1}\left\{\frac{2i-1}{2N} < \gamma^*_{n,m} \leq \frac{2i+1}{2N}\right\} \\
&\quad + \sum_{m=m_n^*+1}^{M_n} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} \sum_{i=0}^{N} \frac{\left(\frac{i}{N} - \gamma^*_{n,m}\right)^2}{\gamma^*_{n,m}} \mathbf{1}\left\{\frac{2i-1}{2N} < \gamma^*_{n,m} \leq \frac{2i+1}{2N}\right\} \\
&\leq \frac{1}{2N}\sum_{m=2}^{m_n^*} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} + \frac{1}{2N}\sum_{m=m_n^*+1}^{M_n} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} \\
&= \frac{1}{2N}\left[\mathrm{tr}\{(\mathbf{P}_{m_n^*} - \mathbf{P}_1)\boldsymbol{\Omega}\} + \boldsymbol{\mu}^\top(\mathbf{P}_{M_n} - \mathbf{P}_{m_n^*})\boldsymbol{\mu}\right]
\end{aligned}$$

$$\leq \frac{1}{2N} R_n^{\mathrm{MS}}(m_n^*),$$

where the first inequality is derived by the fact that when $\frac{2i-1}{2N} < \gamma_{n,m}^* \leq \frac{2i+1}{2N}$ for $i = 0, \ldots, N$,

$$\frac{\left(\frac{i}{N} - \gamma_{n,m}^*\right)^2}{1 - \gamma_{n,m}^*} \leq \frac{1}{2N} \quad \text{and} \quad \frac{\left(\frac{i}{N} - \gamma_{n,m}^*\right)^2}{\gamma_{n,m}^*} \leq \frac{1}{2N},$$

which can be easily verified. Therefore, $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \leq \frac{1}{2N} R_n^{\mathrm{MS}}(m_n^*)$.

When Conditions M2 and A1 hold, our task is to prove that $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$ has the same order as $R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$. We consider two scenarios: $M_n \geq m_n(\frac{2N-1}{2N})$ and $M_n < m_n(\frac{2N-1}{2N})$ but $M_n/m_n^{**} \geq \underline{c}$ for any sufficiently large n.

First, consider $M_n \geq m_n(\frac{2N-1}{2N})$. Define $t_n^N = \min\{t \in \mathbb{N} \colon \lfloor kt \rfloor \geq m_n(\frac{2N-1}{2N}) + 1\}$. Then it follows from Theorem 2 and Peng and Yang (2022) that $\lim_{n\to\infty} t_n^N = \infty$ and $\lfloor kt_n^N \rfloor \sim m_n(\frac{2N-1}{2N})$, respectively. Using the same arguments as that in (S.3) and (S.4), we have

$$\begin{aligned}
&\sum_{m=2}^{m_n^*} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\}(1 - \gamma_{n,m}^*)\mathbf{1}\{\gamma_{n,m}^* > 1 - 1/(2N)\} \\
&= \sum_{m=2}^{m_n(\frac{2N-1}{2N})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\} - \sum_{m=2}^{m_n(\frac{2N-1}{2N})} \frac{\mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\}}{1 + 1/(n\theta_{n,m})} \\
&\geq \mathrm{tr}\{(\mathbf{P}_{m_n(\frac{2N-1}{2N})} - \mathbf{P}_{t_n^N})\mathbf{\Omega}\} - \frac{1}{1 + 1/(n\theta_{n,t_n^N})} \mathrm{tr}\{(\mathbf{P}_{\lfloor kt_n^N \rfloor} - \mathbf{P}_{t_n^N})\mathbf{\Omega}\} \\
&\geq \frac{1}{1 + \frac{\delta}{2N-1}} \mathrm{tr}\left[\left\{\left(1 + \frac{\delta}{2N-1}\right)\mathbf{P}_{m_n(\frac{2N-1}{2N})} - \mathbf{P}_{\lfloor kt_n^N \rfloor} - \frac{\delta}{2N-1}\mathbf{P}_{t_n}\right\}\mathbf{\Omega}\right] \\
&\geq \frac{c_1}{1 + \frac{\delta}{2N-1}} \left\{\left(1 + \frac{\delta}{2N-1}\right)\nu_{m_n(\frac{2N-1}{2N})} - \nu_{\lfloor kt_n^N \rfloor} - \frac{\delta}{2N-1}\nu_{t_n^N}\right\} \\
&\geq \frac{c_1}{1 + \frac{\delta}{2N-1}} (\nu_{m_n(\frac{2N-1}{2N})} - \nu_{\lfloor kt_n^N \rfloor}) + \frac{c_1\delta}{2N - 1 + \delta}\left\{m_n\left(\frac{2N-1}{2N}\right) - t_n^N\right\} \\
&\sim \frac{c_1(k-1)\delta}{k(2N-1+\delta)} m_n\left(\frac{2N-1}{2N}\right) \asymp m_n^*,
\end{aligned} \tag{S.18}$$

where the second inequality is derived by the fact

$$\frac{1}{1 + 1/(n\theta_{n,t_n^N})} \leq \frac{1}{1 + \delta/(n\theta_{n,\lfloor kt_n^N \rfloor})} \leq \frac{1}{1 + \delta/(n\theta_{n,m_n(\frac{2N-1}{2N})+1})} \leq \frac{1}{1 + \delta/(2N-1)},$$

and the last line is due to $\nu_{m_n(\frac{2N-1}{2N})} \sim \nu_{\lfloor kt_n^N \rfloor}$, $t_n^N \sim m_n(\frac{2N-1}{2N})/k$, and Lemma S.2(ii). Since

$$\frac{1}{2N} R_n^{\mathrm{MS}}(m_n^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \geq \sum_{m=2}^{m_n^*} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\}(1 - \gamma_{n,m}^*)\mathbf{1}\{\gamma_{n,m}^* > 1 - 1/(2N)\},$$

using (S.18) and $\mathrm{tr}(\mathbf{P}_{m_n^*}\mathbf{\Omega}) \asymp R_n^{\mathrm{MS}}(m_n^*) \asymp R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$, we have $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \asymp R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$.

Next, consider $M_n < m_n(\frac{2N-1}{2N})$ but $M_n/m_n^{**} \geq \underline{c}$. Using (S.11) and the similar arguments in (S.18), we can also prove $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \asymp R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$. This completes the proof of Theorem 6 under Conditions M2 and A1.

When Condition M1 or Conditions M2 and A2 hold, $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*) = o\{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)\}$ directly follows from Theorems 3–4 and the fact $R_n^{\mathrm{MS}}(m_n^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$.

S.2 Proof of the Results in Examples 5.1–5.2

Using the expression of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)$ in Lemma S.1, we have that for any sufficiently large n,

$$\begin{aligned}\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) = \frac{\sigma^2}{n} &+ \sum_{i=i_{n,N}+1}^{N}\sum_{m=m_n(\frac{2i+1}{2N})+1}^{m_n(\frac{2i-1}{2N})}\left\{\frac{\sigma^2}{n}\left(\frac{i}{N}\right)^2 + \left(1-\frac{i}{N}\right)^2\beta_m^2\right\} \\ &+ \sum_{m=m_n(\frac{2i_{n,N}+1}{2N})+1}^{M_n}\left\{\frac{\sigma^2}{n}\left(\frac{i_{n,N}}{N}\right)^2 + \left(1-\frac{i_{n,N}}{N}\right)^2\beta_m^2\right\} + \sum_{m=M_n+1}^{p_n}\beta_m^2.\end{aligned}$$

Proof of the results in Example 5.1: When $\beta_m = m^{-\alpha}$ for $\alpha > 1/2$, we have $m_n(\frac{2i+1}{2N}) \sim (\frac{2N}{2i+1} - 1)^{\frac{1}{2\alpha}}(\frac{n}{\sigma^2})^{\frac{1}{2\alpha}}$ for $i = i_{n,N}, \dots, N-1$ and $m_n^{**} \sim (\frac{n}{\sigma^2})^{\frac{1}{2\alpha}}$. When $M_n \equiv M$ is fixed as $n \to \infty$, $i_{n,N} = N$ for any sufficiently large n. Therefore,

$$\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) = \frac{M\sigma^2}{n} + \sum_{m=M+1}^{p_n} m^{-2\alpha} \sim \sum_{m=M+1}^{\infty} m^{-2\alpha}.$$

When $M_n \to \infty$ as $n \to \infty$, the optimal risk of MA restricted to $\mathcal{W}_n(N)$ satisfies

$$\begin{aligned}\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) &\sim \frac{\sigma^2}{n} + \sum_{i=i_{n,N}+1}^{N}\int_{m_n(\frac{2i+1}{2N})}^{m_n(\frac{2i-1}{2N})}\left\{\frac{\sigma^2}{n}\left(\frac{i}{N}\right)^2 + \left(1-\frac{i}{N}\right)^2 x^{-2\alpha}\right\}dx \\ &\quad + \int_{m_n(\frac{2i_{n,N}+1}{2N})}^{M_n}\left\{\frac{\sigma^2}{n}\left(\frac{i_{n,N}}{N}\right)^2 + \left(1-\frac{i_{n,N}}{N}\right)^2 x^{-2\alpha}\right\}dx + \int_{M_n}^{p_n} x^{-2\alpha}\,dx \\ &\equiv \frac{\sigma^2}{n} + \Pi_{n1} + \Pi_{n2} + \frac{1}{2\alpha-1}(M_n^{-2\alpha+1} - p_n^{-2\alpha+1}). \qquad \text{(S.1)}\end{aligned}$$

Since $m_n^{**} \sim (\frac{n}{\sigma^2})^{\frac{1}{2\alpha}}$, it is easy to see that $i_{n,N} \sim i_{n,N}^* \equiv \left\lceil \frac{N}{1+(\frac{M_n}{m_n^{**}})^{2\alpha}} - \frac{1}{2}\right\rceil$. We first simplify

Π_{n1} as follows:

$$\Pi_{n1} = \frac{\sigma^2}{n}\sum_{i=i_{n,N}+1}^{N}\left(\frac{i}{N}\right)^2\left\{m_n\left(\frac{2i-1}{2N}\right)-m_n\left(\frac{2i+1}{2N}\right)\right\}$$
$$-\frac{1}{2\alpha-1}\sum_{i=i_{n,N}+1}^{N}\left(1-\frac{i}{N}\right)^2\left\{m_n\left(\frac{2i-1}{2N}\right)^{1-2\alpha}-m_n\left(\frac{2i+1}{2N}\right)^{1-2\alpha}\right\}$$
$$=\frac{\sigma^2}{n}\frac{2}{N}\sum_{i=i_{n,N}}^{N-1}\left(\frac{2i+1}{2N}\right)m_n\left(\frac{2i+1}{2N}\right)+\frac{1}{2\alpha-1}\frac{2}{N}\sum_{i=i_{n,N}}^{N-1}\left(1-\frac{2i+1}{2N}\right)m_n\left(\frac{2i+1}{2N}\right)^{1-2\alpha}$$
$$-\frac{\sigma^2}{n}+\frac{\sigma^2}{n}\left(\frac{i_{n,N}}{N}\right)^2 m_n\left(\frac{2i_{n,N}+1}{2N}\right)-\frac{1}{2\alpha-1}\left(1-\frac{i_{n,N}}{2N}\right)^2 m_n\left(\frac{2i_{n,N}+1}{2N}\right)^{1-2\alpha}$$
$$\sim\frac{2\alpha}{2\alpha-1}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}-1}\frac{2}{N}\sum_{i=i^*_{n,N}}^{N-1}\left(\frac{2i+1}{2N}\right)^{1-\frac{1}{2\alpha}}\left(1-\frac{2i+1}{2N}\right)^{\frac{1}{2\alpha}}-\frac{\sigma^2}{n}$$
$$+\frac{\sigma^2}{n}\left(\frac{i_{n,N}}{N}\right)^2 m_n\left(\frac{2i_{n,N}+1}{2N}\right)-\frac{1}{2\alpha-1}\left(1-\frac{i_{n,N}}{2N}\right)^2 m_n\left(\frac{2i_{n,N}+1}{2N}\right)^{1-2\alpha}. \tag{S.2}$$

Next, we simplify Π_{n2} as follows:

$$\Pi_{n2} = \frac{\sigma^2}{n}\left(\frac{i_{n,N}}{N}\right)^2\left\{M_n-m_n\left(\frac{2i_{n,N}+1}{2N}\right)\right\}$$
$$-\frac{1}{2\alpha-1}\left(1-\frac{i_{n,N}}{N}\right)^2\left\{M_n^{1-2\alpha}-m_n\left(\frac{2i_{n,N}+1}{2N}\right)^{1-2\alpha}\right\}$$
$$\sim\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}-1}\left(\frac{i^*_{n,N}}{N}\right)^2\frac{M_n}{m_n^{**}}-\frac{1}{2\alpha-1}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}-1}\left(1-\frac{i^*_{n,N}}{N}\right)^2\left(\frac{M_n}{m_n^{**}}\right)^{1-2\alpha}$$
$$-\frac{\sigma^2}{n}\left(\frac{i_{n,N}}{N}\right)^2 m_n\left(\frac{2i_{n,N}+1}{2N}\right)+\frac{1}{2\alpha-1}\left(1-\frac{i_{n,N}}{2N}\right)^2 m_n\left(\frac{2i_{n,N}+1}{2N}\right)^{1-2\alpha}. \tag{S.3}$$

Combining (S.1), (S.2), and (S.3), we have that when $M_n\to\infty$ as $n\to\infty$,

$$\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,N})\sim\frac{2\alpha}{2\alpha-1}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}-1}\psi_{n,N}+\frac{1}{2\alpha-1}(M_n^{-2\alpha+1}-p_n^{-2\alpha+1}), \tag{S.4}$$

where

$$\psi_{n,N}=\frac{2}{N}\sum_{i=i^*_{n,N}}^{N-1}\left(\frac{2i+1}{2N}\right)^{1-\frac{1}{2\alpha}}\left(1-\frac{2i+1}{2N}\right)^{\frac{1}{2\alpha}}+\frac{2\alpha-1}{2\alpha}\left(\frac{i^*_{n,N}}{N}\right)^2\frac{M_n}{m_n^{**}}-\frac{1}{2\alpha}\left(1-\frac{i^*_{n,N}}{N}\right)^2\left(\frac{M_n}{m_n^{**}}\right)^{1-2\alpha}.$$

When $M_n\to\infty$ as $n\to\infty$, it is shown in Peng and Yang (2022) that the optimal risk of MA with the weight set $\mathcal{W}_n$ satisfies

$$\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}^*_n)\sim\frac{1}{2\alpha}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}-1}\left\{\frac{\pi}{\sin(\frac{\pi}{2\alpha})}-B\left(\frac{1}{1+\left(\frac{M_n}{m_n^{**}}\right)^{2\alpha}};1-\frac{1}{2\alpha},\frac{1}{2\alpha}\right)\right\}+\frac{1}{2\alpha-1}(M_n^{-2\alpha+1}-p_n^{-2\alpha+1}). \tag{S.5}$$

When $M_n \equiv M$ is fixed as $n \to \infty$,

$$\frac{1}{n} R_n^{\mathrm{MA}}(\mathbf{w}_n^*) = \frac{\sigma^2}{n} + \sum_{m=2}^{M} \frac{1}{\frac{n}{\sigma^2} + m^{2\alpha}} + \sum_{m=M+1}^{p_n} m^{-2\alpha} \sim \sum_{m=M+1}^{\infty} m^{-2\alpha}.$$

Therefore, we consider different conditions on M_n as follows.

(i) When $M_n \equiv M$ is fixed as $n \to \infty$,

$$\frac{1}{n} R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) \sim \frac{1}{n} R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \sim \sum_{m=M+1}^{\infty} m^{-2\alpha}.$$

(ii) When $M_n \to \infty$ but $M_n/m_n^{**} \to 0$ as $n \to \infty$, we have $i_{n,N}^* = N$ for any sufficiently large n, and thus $\psi_{n,N} = o(1)$, which, along with the fact that $B(1; 1 - \frac{1}{2\alpha}, \frac{1}{2\alpha}) = \frac{\pi}{\sin(\frac{\pi}{2\alpha})}$, yields that

$$\frac{1}{n} R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) \sim \frac{1}{n} R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \sim \frac{1}{2\alpha - 1}(M_n^{-2\alpha+1} - p_n^{-2\alpha+1}) \sim \frac{M_n^{-2\alpha+1}}{2\alpha - 1}.$$

(iii) When $M_n/m_n^{**} \geq \underline{c}$ for some $\underline{c} > 0$, let us find the lower bound of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$. If $M_n \geq m_n(\frac{2N-1}{2N})$, note that

$$\begin{aligned}
&\sum_{m=2}^{m_n^*} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\}(1 - \gamma_{n,m}^*)\mathbf{1}\{\gamma_{n,m}^* > 1 - 1/(2N)\} \\
&= \sum_{m=2}^{m_n(\frac{2N-1}{2N})} \sigma^2 \left(1 - \frac{\beta_m^2}{\beta_m^2 + \sigma^2/n}\right) \geq \sum_{m=\lfloor m_n(\frac{2N-1}{2N})/2 \rfloor}^{m_n(\frac{2N-1}{2N})} \frac{\sigma^4/n}{m^{-2\alpha} + \sigma^2/n} \\
&\geq \left\lfloor \frac{1}{2} m_n \left(\frac{2N-1}{2N}\right) \right\rfloor \frac{\sigma^4/n}{\lfloor m_n(\frac{2N-1}{2N})/2 \rfloor^{-2\alpha} + \sigma^2/n} \sim \frac{(2N-1)^{-\frac{1}{2\alpha}} \sigma^2}{2^{2\alpha+1}(2N-1) + 2} \left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}}.
\end{aligned}$$

If $m_n(\frac{2N-1}{2N}) > M_n \geq \underline{c} m_n^{**}$, we also have

$$\begin{aligned}
&\sum_{m=2}^{m_n^*} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\mathbf{\Omega}\}(1 - \gamma_{n,m}^*)\mathbf{1}\{\gamma_{n,m}^* > 1 - 1/(2N)\} \\
&\geq \lfloor \underline{c} m_n^{**}/2 \rfloor \frac{\sigma^4/n}{\lfloor \underline{c} m_n^{**}/2 \rfloor^{-2\alpha} + \sigma^2/n} \sim \frac{\underline{c}\sigma^2}{2^{2\alpha+1}\underline{c}^{-2\alpha} + 2} \left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}}.
\end{aligned}$$

As a result, $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$ can be lower bounded by $\frac{\varpi\sigma^2}{2^{2\alpha+1}\varpi^{-2\alpha}+2}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}}$, where $\varpi = \min\{\underline{c}, (2N-1)^{-\frac{1}{2\alpha}}\}$. Moreover, if $\lim_{n\to\infty} M_n/m_n^{**} = \kappa \in (0, \infty]$ and $M_n = o(p_n)$ are satisfied, it follows from (S.4) and (S.5) that

$$\lim_{n\to\infty} \frac{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}{R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)} = \frac{1}{\psi_N^* + \frac{\kappa^{-2\alpha+1}}{2\alpha}} \left[\frac{2\alpha - 1}{4\alpha^2}\left\{\frac{\pi}{\sin(\frac{\pi}{2\alpha})} - B\left(\frac{1}{1+\kappa^{2\alpha}}; 1 - \frac{1}{2\alpha}, \frac{1}{2\alpha}\right)\right\} + \frac{\kappa^{-2\alpha+1}}{2\alpha}\right],$$

where

$$\psi_N^* = \frac{2}{N}\sum_{i=i_N^*}^{N-1}\left(\frac{2i+1}{2N}\right)^{1-\frac{1}{2\alpha}}\left(1-\frac{2i+1}{2N}\right)^{\frac{1}{2\alpha}}+\frac{2\alpha-1}{2\alpha}\left(\frac{i_N^*}{N}\right)^2\kappa-\frac{1}{2\alpha}\left(1-\frac{i_N^*}{N}\right)^2\kappa^{1-2\alpha} \quad \text{(S.6)}$$

and $i_N^* = \lceil \frac{N}{1+\kappa^{2\alpha}} - \frac{1}{2} \rceil$. It is easy to see that $\{\psi_N^*\}_{N=1}^{\infty}$ is a strictly decreasing sequence with $\psi_1^* = 1 - \frac{\kappa^{-2\alpha+1}}{2\alpha}$. Moreover, we can prove that

$$\begin{aligned}\lim_{N\to\infty}\psi_N^* &= 2\int_{\frac{1}{1+\kappa^{2\alpha}}}^1 t^{1-\frac{1}{2\alpha}}(1-t)^{\frac{1}{2\alpha}}\,dt + \frac{2\alpha-1}{2\alpha}\frac{\kappa}{(1+\kappa^{2\alpha})^2} - \frac{1}{2\alpha}\frac{\kappa^{1+2\alpha}}{(1+\kappa^{2\alpha})^2}\\ &= \frac{2\alpha-1}{4\alpha^2}\int_{\frac{1}{1+\kappa^{2\alpha}}}^1 t^{-\frac{1}{2\alpha}}(1-t)^{\frac{1}{2\alpha}-1}\,dt\\ &= \frac{2\alpha-1}{4\alpha^2}\left\{\frac{\pi}{\sin(\frac{\pi}{2\alpha})} - B\left(\frac{1}{1+\kappa^{2\alpha}};1-\frac{1}{2\alpha},\frac{1}{2\alpha}\right)\right\},\end{aligned}$$

where the last equality follows from the fact that $B(1;1-\frac{1}{2\alpha},\frac{1}{2\alpha}) = \frac{\pi}{\sin(\frac{\pi}{2\alpha})}$. Therefore, for any fixed $N \geq 1$,

$$\lim_{n\to\infty}\frac{R_n^{\mathrm{MA}}(\mathbf{w}_n^*)}{R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)} < 1.$$

Proof of the results in Example 5.2: When $\beta_m = \exp(-cm)$ for $c > 0$, we have $m_n(\frac{2i+1}{2N}) \sim \frac{1}{2c}\log\left(\frac{n}{\sigma^2}\right)$ for $i = i_{n,N},\ldots,N-1$ and $m_n^{**} \sim \frac{1}{2c}\log\left(\frac{n}{\sigma^2}\right)$. The optimal risk of MA satisfies

$$\begin{aligned}\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_n^*) &= \frac{\sigma^2}{n} + \sum_{m=2}^{M_n}\frac{1}{\frac{n}{\sigma^2}+\exp(2cm)} + \sum_{m=M_n+1}^{p_n}\exp(-2cm)\\ &\sim \sum_{m=1}^{M_n}\frac{1}{\frac{n}{\sigma^2}+\exp(2cm)} + \frac{\exp(-2cM_n)-\exp(-2cp_n)}{\exp(2c)-1}. \end{aligned}\quad \text{(S.7)}$$

We consider different conditions on M_n as follows.

(i) When $\limsup_{n\to\infty} M_n/m_n^{**} < 1$, we have $M_n < m_n^{**}$ for any sufficiently large n. Thus,

$$\frac{1}{n}R_n^{\mathrm{MS}}(m_n^*) = \frac{M_n\sigma^2}{n} + \sum_{m=M_n+1}^{p_n}\exp(-2cm) = \frac{M_n\sigma^2}{n} + \frac{\exp(-2cM_n)-\exp(-2cp_n)}{\exp(2c)-1}. \quad \text{(S.8)}$$

By $2cm_n^{**} \sim \log\left(\frac{n}{\sigma^2}\right)$ and $\lim_{n\to\infty}\log(M_n)/m^{**} = 0$, we observe that

$$\limsup_{n\to\infty}\log\left\{\frac{M_n\sigma^2/n}{\exp(-2cM_n)}\right\}\Big/(2cm_n^{**}) = \limsup_{n\to\infty}\frac{\log M_n - \log(\frac{n}{\sigma^2}) + 2cM_n}{2cm_n^{**}} \leq -1+\limsup_{n\to\infty}\frac{M_n}{m_n^{**}} < 0,$$

which implies that $\frac{M_n\sigma^2/n}{\exp(-2cM_n)} \to 0$ as $n\to\infty$. Moreover, as $n\to\infty$,

$$\frac{\exp(-2cp_n)}{\exp(-2cM_n)} = \exp\{-2c(p_n - M_n)\} \leq \exp\{-2c(m_n^{**}-M_n)\} \to 0.$$

Therefore, we have $\frac{1}{n}R_n^{\mathrm{MS}}(m_n^*) \sim \frac{\exp(-2cM_n)}{\exp(2c)-1}$. Since $\sum_{m=1}^{M_n}\{\frac{n}{\sigma^2}+\exp(2cm)\}^{-1} \leq \frac{\sigma^2}{n}M_n$, from (S.7), we have $\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \sim \frac{\exp(-2cM_n)}{\exp(2c)-1}$. Therefore,

$$\frac{1}{n}R_n^{\mathrm{MS}}(m_n^*) \sim \frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \sim \frac{\exp(-2cM_n)}{\exp(2c)-1}.$$

(ii) When $M_n \geq m_n^{**}$ for any sufficiently large n, note that as $n \to \infty$,

$$\frac{\exp(-2cm_n^{**})}{m_n^{**}\sigma^2/n} \leq \frac{\exp(-2cm_n^{**})}{m_n^{**}\exp\{-2c(m_n^{**}+1)\}} = \frac{\exp(2c)}{m_n^{**}} \to 0, \tag{S.9}$$

where the inequality is due to $\sigma^2/n \geq \exp\{-2c(m_n^{**}+1)\}$ derived from (A.2). Therefore, we have

$$\frac{1}{n}R_n^{\mathrm{MS}}(m_n^*) = \frac{m_n^{**}\sigma^2}{n} + \frac{\exp(-2cm_n^{**}) - \exp(-2cp_n)}{\exp(2c)-1} \sim \frac{m_n^{**}\sigma^2}{n}.$$

Next, we investigate $R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$. From (S.7),

$$\begin{aligned}\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_n^*) &\sim \sum_{m=1}^{M_n}\frac{1}{\frac{n}{\sigma^2}+\exp(2cm)} + \sum_{m=M_n+1}^{p_n}\exp(-2cm) \\ &= \sum_{m=1}^{m_n^{**}}\frac{1}{\frac{n}{\sigma^2}+\exp(2cm)} + \sum_{m=m_n^{**}+1}^{M_n}\frac{1}{\frac{n}{\sigma^2}+\exp(2cm)} + \sum_{m=M_n+1}^{p_n}\exp(-2cm). \end{aligned} \tag{S.10}$$

For the first term of (S.10), it is easy to obtain

$$\begin{aligned}\sum_{m=1}^{m_n^{**}}\frac{1}{\frac{n}{\sigma^2}+\exp(2cm)} &\sim \int_0^{m_n^{**}}\frac{1}{\frac{n}{\sigma^2}+\exp(2cx)}\,dx \\ &= \frac{m_n^{**}\sigma^2}{n} - \frac{1}{2c}\frac{\sigma^2}{n}\log\left\{\frac{1+\frac{\sigma^2}{n}\exp(2cm_n^{**})}{1+\frac{\sigma^2}{n}}\right\} \sim \frac{m_n^{**}\sigma^2}{n}, \end{aligned} \tag{S.11}$$

where the last "$\sim$" is due to $\frac{\sigma^2}{n}\exp(2cm_n^{**}) < 1$ derived from (A.2). For the last two terms of (S.10), using (S.9), we have

$$\begin{aligned}&\sum_{m=m_n^{**}+1}^{M_n}\frac{1}{\frac{n}{\sigma^2}+\exp(2cm)} + \sum_{m=M_n+1}^{p_n}\exp(-2cm) \\ &\leq \sum_{m=m_n^{**}+1}^{p_n}\exp(-2cm) = \frac{\exp(-2cm_n^{**}) - \exp(-2cp_n)}{\exp(2c)-1} = o\left(\frac{m_n^{**}\sigma^2}{n}\right), \end{aligned}$$

which, along with (S.10) and (S.11), yields that $\frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \sim \frac{m_n^{**}\sigma^2}{n}$. Therefore,

$$\frac{1}{n}R_n^{\mathrm{MS}}(m_n^*) \sim \frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \sim \frac{m_n^{**}\sigma^2}{n} \sim \frac{1}{2c}\frac{\sigma^2}{n}\log\left(\frac{n}{\sigma^2}\right).$$

(iii) When $M_n < m_n^{**}$ for any sufficiently large n but $\lim_{n\to\infty} M_n/m_n^{**} = 1$, by using the same arguments in (S.11), we can show that $\sum_{m=1}^{M_n}\{\frac{n}{\sigma^2}+\exp(2cm)\}^{-1} \sim \frac{M_n\sigma^2}{n}$, which, along with (S.7) and (S.8), yields that

$$\frac{1}{n}R_n^{\mathrm{MS}}(m_n^*) \sim \frac{1}{n}R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \sim \frac{1}{2c}\frac{\sigma^2}{n}\log\left(\frac{n}{\sigma^2}\right) + \frac{\exp(-2cM_n)-\exp(-2cp_n)}{\exp(2c)-1}.$$

Combining results (i)–(iii) and the fact $R_n^{\mathrm{MS}}(m_n^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$, we obtain the results of Example 5.2.

S.3 A Comparison of MA Techniques with Nested Discrete Weight Sets

S.3.1 Question Q5

In addition to the proposed four questions in Section 2, another natural question is to compare the optimal risks of MS and MA restricted to $\mathcal{W}_n(N)$. Note that MA restricted to $\mathcal{W}_n(1)$ reduces to MS. Therefore, we can investigate a more general problem that compares the optimal risks of MA techniques with weights belonging to two nested discrete weight sets $\mathcal{W}_n(d)$ and $\mathcal{W}_n(dN)$, where $d \geq 1$ and $N \geq 2$ are fixed integers. Since $\mathcal{W}_n(d)$ is a subset of $\mathcal{W}_n(dN)$, we have $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)$. However, it remains unclear whether expanding the discrete weight set for MA leads to a significant improvement in risk. Thus, the following key question is proposed:

Q5. Is $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)$ a substantial reduction relative to $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)$ or actually negligible? If both can happen, when is $\mathbf{w}_{n,dN}^*$ substantially better than $\mathbf{w}_{n,d}^*$?

S.3.2 An Answer to Question Q5

We first consider two conditions for the number of candidate models M_n as follows:

(i) $\theta_{n,M_n} > (2dN-1)/n$ for sufficiently large n a.s.;

(ii) $M_n \geq m_n^{**}$ for sufficiently large n a.s.

These two conditions are slightly different from Conditions M1 and M2. The first condition (i) restricts M_n not to be too large. Under Condition A1 or A2, if $\lim_{n\to\infty} M_n/m_n^{**} = 0$, then there exist $k > 1$, $\eta \in (0,1)$, and a positive integer s such that on $\mathcal{F}$,

$$\theta_{n,M_n} \geq \theta_{n,\lfloor k^{-s} m_n^{**} \rfloor} \geq \eta^{-s}\theta_{n,m_n^{**}} > \eta^{-s}/n \geq (2dN-1)/n,$$

where the third inequality is due to (A.2) in the Appendix. Therefore, under some mild conditions on $\theta_{n,m}$ (e.g., Condition A1 or A2), the condition (i) is weaker than Condition M1. The second condition (ii) is stronger than Condition M2, which is considered by Peng and Yang (2022).

Next, we make a new condition on the slowly decaying order of $\{\theta_{n,m}\}_{m=1}^{d_n}$ as follows.

Condition C1 (Slowly Decaying $\{\theta_{n,m}\}_{m=1}^{d_n}$). There exist constants $k > 1$, $\frac{2d-1}{2dN-1} < \delta \leq \eta < 1$ with $k\eta < 1$, and $K > 0$ such that for every integer sequence $\{l_n\}$ satisfied $\lim_{n\to\infty} l_n = \infty$,

$$\delta \leq \theta_{n,\lfloor k l_n \rfloor}/\theta_{n,l_n} \leq \eta$$

holds for any $n \geq K$ a.s.

Condition C1 is stronger than Condition A1 since Condition C1 additionally requires that $\delta > \frac{2d-1}{2dN-1}$, which restricts δ to not close to 0. Note that when $d = 1$, Condition C1 restricts $\delta > \frac{1}{2N-1}$. Condition C1 is still satisfied for the polynomial decay case, e.g., $\theta_{n,l_n} \sim l_n^{-2\alpha}$, $\alpha > 1/2$ or slightly more generally for $\theta^*_{n,l_n} \sim l_n^{-2\alpha}(\log l_n)^{\beta}$, $\alpha > 1/2, \beta \in \mathbb{R}$, where $\{l_n\}$ is an integer sequence satisfied $\lim_{n\to\infty} l_n = \infty$.

Now, we turn our attention to answer Question Q5 in the following theorem.

Theorem 7 (**Answer to Question Q5**). *Suppose that Assumptions 1–6 hold. Then, for sufficiently large n,*

(i) when $\theta_{n,M_n} > (2dN-1)/n$, we have $R_n^{\mathrm{MS}}(m_n^*) = R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,2}) = \cdots = R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,dN})$ a.s.;

(ii) when $M_n \geq m_n^{**}$, under Condition C1, we have

$$R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,d}) - R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,dN}) \asymp R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,d}) \text{ a.s.};$$

and under Condition A2, we have

$$R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,d}) - R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,dN}) = o\{R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,d})\} \text{ a.s.}$$

Theorem 7(i) implies that when the number of candidate models M_n satisfies $\theta_{n,M_n} > (2dN-1)/n$, the optimal risk of MA remains unchanged for sufficiently large n when the discrete weight set is expanded from $\mathcal{W}_n(1)$ to $\mathcal{W}_n(dN)$. Theorem 7(ii) implies that when M_n is large enough and $\theta_{n,m}$ decays slowly in m, expanding the discrete weight set of MA can bring in a substantial reduction in risk. When M_n is large enough and $\theta_{n,m}$ decays fast in m, the risk reduction of MA by expanding the discrete weight set is asymptotically negligible.

Next, we consider the case of $d = 1$, i.e., we compare the optimal risks of MS and MA restricted to the discrete set $\mathcal{W}_n(N)$, where $N \geq 2$ is a fixed integer. From Theorem 7, we have the following corollary on a comparison of $R_n^{\mathrm{MS}}(m_n^*)$ and $R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)$.

Corollary 4. *Suppose that Assumptions 1–6 hold. Then, for sufficiently large n,*

(i) when $\theta_{n,M_n} > (2N-1)/n$, we have $R_n^{\mathrm{MS}}(m_n^*) = R_n^{\mathrm{MA}}(\mathbf{w}_{n,2}^*) = \cdots = R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)$ a.s.;

(ii) when $M_n \geq m_n^{**}$, under Condition C1 with $d = 1$, $R_n^{\mathrm{MS}}(m_n^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) \asymp R_n^{\mathrm{MS}}(m_n^*)$ a.s.; and under Condition A2, $R_n^{\mathrm{MS}}(m_n^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*) = o\{R_n^{\mathrm{MS}}(m_n^*)\}$ a.s.

Example 5.1 (Continued). In the setting of Example 5.1, we consider $M_n \geq m_n^{**}$ for sufficiently large n and any fixed $d \geq 1$ and $N \geq 2$. By a simple calculation, $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)$ is lower bounded by $\frac{2\sigma^2}{3dN}\{(2d-1)^{-\frac{1}{2\alpha}} - (\frac{4}{3}dN-1)^{-\frac{1}{2\alpha}}\}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}}$. Moreover, if $\lim_{n\to\infty} M_n/m_n^* = \kappa$, $\kappa \in [1,\infty]$ and $M_n = o(p_n)$, we have

$$\lim_{n\to\infty} \frac{R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)}{R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)} = \frac{\psi_{dN}^* + \frac{\kappa^{-2\alpha+1}}{2\alpha}}{\psi_d^* + \frac{\kappa^{-2\alpha+1}}{2\alpha}} < 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)}{R_n^{\mathrm{MS}}(m_n^*)} = \psi_N^* + \frac{\kappa^{-2\alpha+1}}{2\alpha} < 1,$$

where ψ_N^* is defined in (S.6), which verifies that $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*) = o\{R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)\}$ in Theorem 7. Figure S.1 plots $\lim_{n\to\infty} R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)/R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)$ against different N or κ, where $\alpha = 0.8$. Specifically,

- Figure S.1(a)–(b) display plots of $\lim_{n\to\infty} R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)/R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)$ against $N \in \{1,\ldots,10\}$ for $d = 1,\ldots,4$, where (a): $\kappa = 0.2$; (b): $\kappa = 1.5$.

- Figure S.1(c)–(d) display plots of $\lim_{n\to\infty} R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)/R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)$ against $\kappa \in (0,4)$, where (c): $d = 1$ and $N = 2,3,8$; (d): $d = 1,2,4$ and $N = 2$.

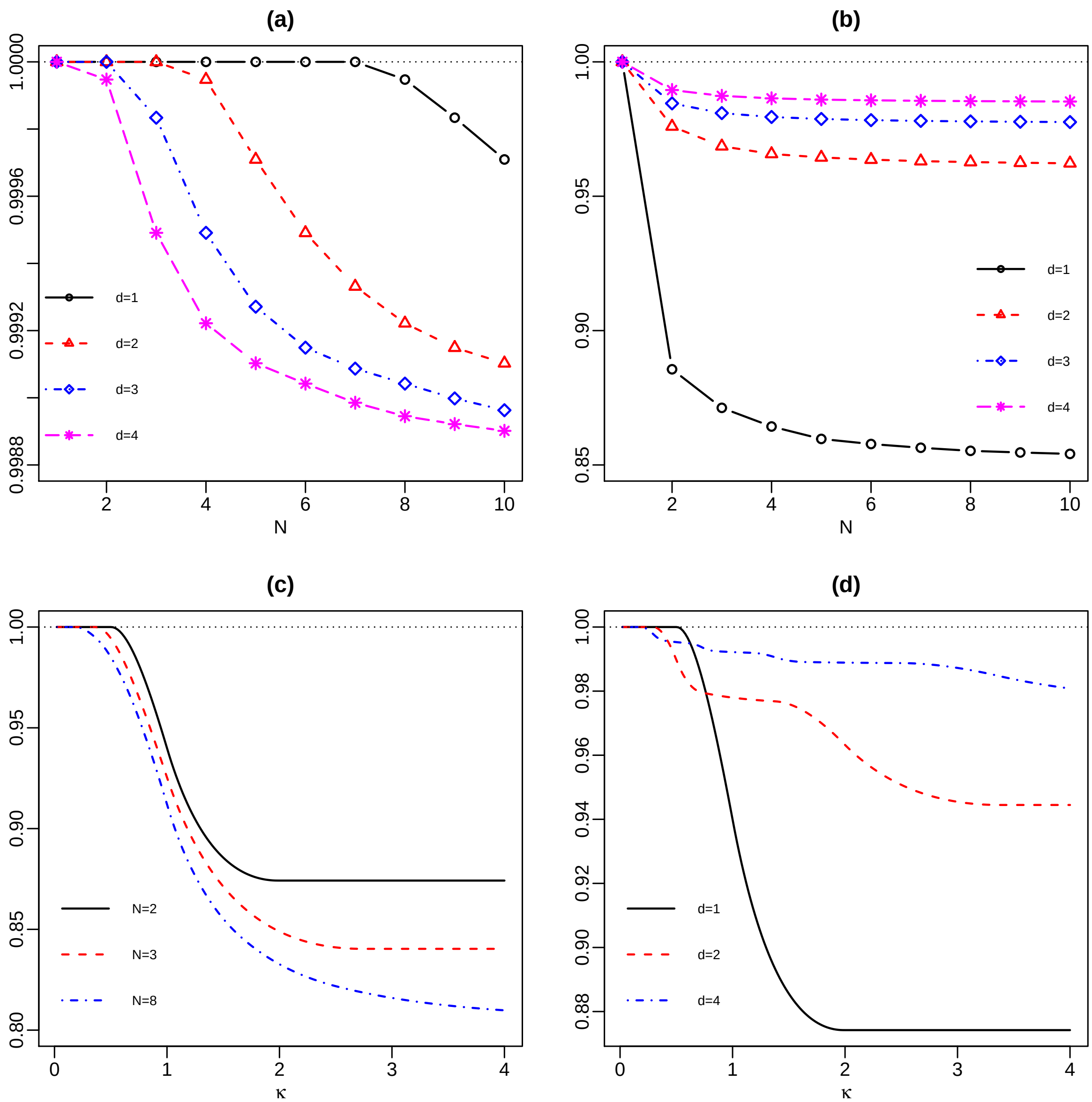


Figure S.1: Numerical illustration for Example 5.1 with $\alpha = 0.8$. (a)–(b): plots of $\lim_{n\to\infty} R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,dN})/R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,d})$ against $N \in \{1, \ldots, 10\}$ for $d = 1, \ldots, 4$, where (a): $\kappa = 0.2$; (b): $\kappa = 1.5$. (c)–(d): plots of $\lim_{n\to\infty} R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,dN})/R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,d})$ against $\kappa \in (0, 4)$, where (c): $d = 1$ and $N = 2, 3, 8$; (d): $d = 1, 2, 4$ and $N = 2$.

Figure S.1(b)–(d) verify that $\lim_{n\to\infty} R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)/R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) < 1$ when $N \geq 2$ and $\kappa \geq 1$. Figure S.1(a) implies that $\lim_{n\to\infty} R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)/R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) < 1$ may not hold for small N when $\kappa < 1$.

S.3.3 Proofs of the Main Results

Proof of Theroem 7. We consider the following two cases (i)–(ii).

(i) $\theta_{n,M_n} > (2dN-1)/n$ for sufficiently large n. By the definition of $m_n(z)$ in (S.13), we have $\theta_{n,M_n} > (2dN-1)/n \geq \theta_{n,m_n(\frac{2dN-1}{2dN})+1}$, which follows that $M_n < m_n(\frac{2dN-1}{2dN}) + 1$ and $\gamma_{n,M_n}^* > 1 - \frac{1}{2dN}$. For any $h = 1, \dots, dN$, it is easy to see that $i_{n,h}$ defined in Lemma S.1 satisfies $i_{n,h} = \lceil h\gamma_{n,M_n}^* - \frac{1}{2} \rceil = h$. Then, from the expression of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,h}^*)$ in Lemma S.1, we obtain

$$R_n^{\mathrm{MA}}(\mathbf{w}_{n,h}^*) = \mathrm{tr}(\mathbf{P}_{M_n}\boldsymbol{\Omega}) + \boldsymbol{\mu}^\top(\mathbf{I}_n - \mathbf{P}_{M_n})\boldsymbol{\mu} = R_n^{\mathrm{MS}}(m_n^*)$$

for sufficiently large n, which leads to $R_n^{\mathrm{MS}}(m_n^*) = R_n^{\mathrm{MA}}(\mathbf{w}_{n,2}^*) = \cdots = R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)$.

(ii) $M_n \geq m_n^{**}$ for sufficiently large n. We first present an expression of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)$ as follows. By (S.15) and the definition of $m_n(z)$ in Lemma S.1, it is easy to rewrite $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$ as

$$\begin{aligned}
&R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \\
&\quad= \sum_{i=r_{n,d}+1}^{2d-1} \sum_{m=m_n(\frac{i+1}{2d})+1}^{m_n(\frac{i}{2d})} \left[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\right] \left(\frac{\lceil i/2 \rceil}{d} - \gamma_{n,m}^*\right)^2 \\
&\qquad+ \sum_{m=m_n(\frac{r_{n,d}+1}{2d})+1}^{M_n} \left[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\right] \left(\frac{\lceil r_{n,d}/2 \rceil}{d} - \gamma_{n,m}^*\right)^2,
\end{aligned}$$

where $r_{n,d} = \lceil 2d\gamma_{n,M_n}^* - 1 \rceil$. Moreover, $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$ can be further rewritten as

$$\begin{aligned}
&R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_n^*) \\
&\quad= \sum_{i=r_{n,dN}+1}^{2dN-1} \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \left[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\right] \left(\frac{\lceil [i/N]/2 \rceil}{d} - \gamma_{n,m}^*\right)^2 \\
&\qquad+ \sum_{m=m_n(\frac{r_{n,dN}+1}{2dN})+1}^{M_n} \left[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\right] \left(\frac{\lceil [r_{n,dN}/N]/2 \rceil}{d} - \gamma_{n,m}^*\right)^2,
\end{aligned}$$

where $[a]$ denotes the integer part of a. Observe that $[r_{n,dN}/N] = r_{n,d}$. Therefore, an expression of $R_n^{\text{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\text{MA}}(\mathbf{w}_{n,dN}^*)$ is as follows

$$\begin{aligned}
&R_n^{\text{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\text{MA}}(\mathbf{w}_{n,dN}^*) = \left\{R_n^{\text{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\text{MA}}(\mathbf{w}_n^*)\right\} - \left\{R_n^{\text{MA}}(\mathbf{w}_{n,dN}^*) - R_n^{\text{MA}}(\mathbf{w}_n^*)\right\} \\
&= \sum_{i=r_{n,dN}+1}^{2dN-2} \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \Bigg\{\Big[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \text{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\Big] \\
&\qquad \times \left(\frac{\lceil [i/N]/2\rceil}{d} - \frac{\lceil i/2\rceil}{dN}\right)\left(\frac{\lceil [i/N]/2\rceil}{d} + \frac{\lceil i/2\rceil}{dN} - 2\gamma_{n,m}^*\right)\Bigg\} \\
&\quad + \sum_{m=m_n(\frac{r_{n,dN}+1}{2dN})+1}^{M_n} \Bigg\{\Big[\boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu} + \text{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\Big] \\
&\qquad \times \left(\frac{\lceil r_{n,d}/2\rceil}{d} - \frac{\lceil r_{n,dN}/2\rceil}{dN}\right)\left(\frac{\lceil r_{n,d}/2\rceil}{d} + \frac{\lceil r_{n,dN}/2\rceil}{dN} - 2\gamma_{n,m}^*\right)\Bigg\}.
\end{aligned}$$

We can easily verify that when $m_n(\frac{i+1}{2dN}) + 1 \le m \le m_n(\frac{i}{2dN})$, $i = r_{n,dN}+1, \ldots, 2dN-2$ or $m_n(\frac{r_{n,dN}+1}{2dN}) + 1 \le m \le M_n$, $i = r_{n,dN}$, we have

$$\left(\frac{\lceil [i/N]/2\rceil}{d} - \frac{\lceil i/2\rceil}{dN}\right)\left(\frac{\lceil [i/N]/2\rceil}{d} + \frac{\lceil i/2\rceil}{dN} - 2\gamma_{n,m}^*\right) \ge 0.$$

By using (S.17), we can further rewrite $R_n^{\text{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\text{MA}}(\mathbf{w}_{n,dN}^*)$ as

$$\begin{aligned}
&R_n^{\text{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\text{MA}}(\mathbf{w}_{n,dN}^*) \\
&= \sum_{m=m_n(\frac{r_{n,dN}+1}{2dN})+1}^{M_n} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\left(\frac{\lceil r_{n,d}/2\rceil}{d} - \frac{\lceil r_{n,dN}/2\rceil}{dN}\right)\left(-2 + \frac{\frac{\lceil r_{n,d}/2\rceil}{d} + \frac{\lceil r_{n,dN}/2\rceil}{dN}}{\gamma_{n,m}^*}\right) \\
&\quad + \sum_{j=r_{n,d}}^{d-1} \sum_{i=\max\{jN, r_{n,dN}+1\}}^{(j+1)N-1} \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \boldsymbol{\mu}^\top(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\mu}\left(\frac{\lceil j/2\rceil}{d} - \frac{\lceil i/2\rceil}{dN}\right)\left(-2 + \frac{\frac{\lceil j/2\rceil}{d} + \frac{\lceil i/2\rceil}{dN}}{\gamma_{n,m}^*}\right) \\
&\quad + \sum_{j=d}^{2d-1} \sum_{i=jN}^{\min\{(j+1)N-1, 2dN-2\}} \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \text{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\}\left(\frac{\lceil j/2\rceil}{d} - \frac{\lceil i/2\rceil}{dN}\right)\left(2 - \frac{2 - \frac{\lceil j/2\rceil}{d} - \frac{\lceil i/2\rceil}{dN}}{1 - \gamma_{n,m}^*}\right).
\end{aligned} \tag{S.1}$$

Next, we examine when Condition C1 holds, $R_n^{\text{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\text{MA}}(\mathbf{w}_{n,dN}^*) \asymp R_n^{\text{MA}}(\mathbf{w}_{n,d}^*)$. From (S.1), we have

$$R_n^{\text{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\text{MA}}(\mathbf{w}_{n,dN}^*)$$

$$
\begin{aligned}
&\geq \sum_{i=2dN-N}^{2dN-2} \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \left(1 - \frac{\lceil i/2 \rceil}{dN}\right) \left(2 - \frac{1 - \frac{\lceil i/2 \rceil}{dN}}{1 - \gamma^*_{n,m}}\right) \\
&= \sum_{m=m_n(\frac{2dN-1}{2dN})+1}^{m_n(\frac{dN-1}{dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \frac{1}{dN} \left(2 - \frac{1/(dN)}{1 - \gamma^*_{n,m}}\right) \\
&\quad + \sum_{i=2dN-N}^{2dN-3} \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \left(1 - \frac{\lceil i/2 \rceil}{dN}\right) \left(2 - \frac{1 - \frac{\lceil i/2 \rceil}{dN}}{1 - \gamma^*_{n,m}}\right) \\
&\geq \sum_{m=m_n(\frac{2dN-1-\vartheta}{2dN})+1}^{m_n(\frac{dN-1}{dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \frac{1}{dN} \left(2 - \frac{1/(dN)}{1 - \gamma^*_{n,m}}\right) \\
&\quad + \sum_{i=2dN-N}^{2dN-3} \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \left(1 - \frac{\lceil i/2 \rceil}{dN}\right) \left(2 - \frac{1 - \frac{\lceil i/2 \rceil}{dN}}{1 - \gamma^*_{n,m}}\right) \\
&\geq \frac{1}{dN} \frac{2\vartheta}{1+\vartheta} \sum_{m=m_n(\frac{2dN-1-\vartheta}{2dN})+1}^{m_n(\frac{dN-1}{dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \qquad \text{(S.2)} \\
&\quad + \sum_{i=2dN-N}^{2dN-3} \left(1 - \frac{\lceil i/2 \rceil}{dN}\right) \left(2 - \frac{1 - \frac{\lceil i/2 \rceil}{dN}}{1 - \frac{i+1}{2dN}}\right) \sum_{m=m_n(\frac{i+1}{2dN})+1}^{m_n(\frac{i}{2dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\},
\end{aligned}
$$

where $\vartheta \in (0, 1)$ is a constant which will be specified later and the last inequality follows from the fact that $\frac{i}{2dN} < \gamma^*_{n,m} \leq \frac{i+1}{2dN}$ when $m_n(\frac{i+1}{2dN}) + 1 \leq m \leq m_n(\frac{i}{2dN})$. It is easy to see that when $2dN - N \leq i \leq 2dN - 3$,

$$
\left(1 - \frac{\lceil i/2 \rceil}{dN}\right) \left(2 - \frac{1 - \frac{\lceil i/2 \rceil}{dN}}{1 - \frac{i+1}{2dN}}\right) = \left(1 - \frac{\lceil (i+1)/2 \rceil}{dN}\right) \frac{1 - \frac{\lceil i/2 \rceil}{dN}}{1 - \frac{i+1}{2dN}} \geq 1 - \frac{\lceil (i+1)/2 \rceil}{dN} \geq \frac{1}{dN},
$$

which, along with (S.2), yields that

$$
\begin{aligned}
&R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,d}) - R_n^{\mathrm{MA}}(\mathbf{w}^*_{n,dN}) \\
&\quad \geq \frac{1}{dN} \frac{2\vartheta}{1+\vartheta} \sum_{m=m_n(\frac{2dN-1-\vartheta}{2dN})+1}^{m_n(\frac{dN-1}{dN})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} + \frac{1}{dN} \sum_{m=m_n(\frac{dN-1}{dN})+1}^{m_n(\frac{2d-1}{2d})} \mathrm{tr}\{(\mathbf{P}_m - \mathbf{P}_{m-1})\boldsymbol{\Omega}\} \\
&\quad \geq \frac{1}{dN} \frac{2\vartheta}{1+\vartheta} \mathrm{tr}\{(\mathbf{P}_{m_n(\frac{2d-1}{2d})} - \mathbf{P}_{m_n(\frac{2dN-1-\vartheta}{2dN})})\boldsymbol{\Omega}\} \\
&\quad \geq \frac{c_1}{dN} \frac{2\vartheta}{1+\vartheta} \left\{ m_n\left(\frac{2d-1}{2d}\right) - m_n\left(\frac{2dN-1-\vartheta}{2dN}\right) \right\}. \qquad \text{(S.3)}
\end{aligned}
$$

Observe that

$$\frac{\frac{2d-1}{2d}/(1-\frac{2d-1}{2d})}{\frac{2dN-1-\vartheta}{2dN}/(1-\frac{2dN-1-\vartheta}{2dN})} = \frac{2d-1}{\frac{2dN}{1+\vartheta}-1} \xrightarrow{\vartheta\to 0} \frac{2d-1}{2dN-1}.$$

Since Condition C1 requires $\delta > \frac{2d-1}{2dN-1}$, we can find a small enough $\vartheta > 0$ such that

$$\delta \geq \frac{\frac{2d-1}{2d}/(1-\frac{2d-1}{2d})}{\frac{2dN-1-\vartheta}{2dN}/(1-\frac{2dN-1-\vartheta}{2dN})}.$$

Thus, by applying Lemma S.2(ii) and Lemma S.3 presented at the end of this section, we have

$$m_n\left(\frac{2d-1}{2d}\right) - m_n\left(\frac{2dN-1-\vartheta}{2dN}\right) \asymp m_n\left(\frac{2d-1}{2d}\right) \asymp m_n^*,$$

which, along with (S.3) and $R_n^{\mathrm{MS}}(m_n^*) \asymp \mathrm{tr}(\mathbf{P}_{m_n^*}\mathbf{\Omega})$, leads to $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*) \asymp R_n^{\mathrm{MS}}(m_n^*) \asymp R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)$. This completes the proof of Theorem 7 under Condition C1.

When Condition A2 holds, $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*) = o\{R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)\}$ directly follows from Theorem 4 and the fact $R_n^{\mathrm{MS}}(m_n^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*) \geq R_n^{\mathrm{MA}}(\mathbf{w}_n^*)$. □

Proof of the Results in Example 5.1 (Continued). First, let us find the lower bound of $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)$. For any fixed $d \geq 1$ and $N \geq 2$, by (S.3) and letting $\vartheta = 1/2$, we have

$$\begin{aligned} R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*) &\geq \frac{2\sigma^2}{3dN}\left\{m_n\left(\frac{2d-1}{2d}\right) - m_n\left(\frac{2dN-3/2}{2dN}\right)\right\} \\ &\sim \frac{2\sigma^2}{3dN}\left\{(2d-1)^{-\frac{1}{2\alpha}} - (4dN/3-1)^{-\frac{1}{2\alpha}}\right\}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}}. \end{aligned}$$

Thus, $R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*) - R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)$ is lower bounded by $\frac{2\sigma^2}{3dN}\{(2d-1)^{-\frac{1}{2\alpha}} - (\frac{4}{3}dN-1)^{-\frac{1}{2\alpha}}\}\left(\frac{n}{\sigma^2}\right)^{\frac{1}{2\alpha}}$. Next, if $\lim_{n\to\infty} M_n/m_n^{**} = \kappa \in [1,\infty]$ and $M_n = o(p_n)$ are satisfied, it follows from (S.4) that

$$\lim_{n\to\infty} \frac{R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)}{R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)} = \frac{\psi_{dN}^* + \frac{\kappa^{-2\alpha+1}}{2\alpha}}{\psi_d^* + \frac{\kappa^{-2\alpha+1}}{2\alpha}} \quad \text{and} \quad \lim_{n\to\infty} \frac{R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)}{R_n^{\mathrm{MS}}(m_n^*)} = \psi_N^* + \frac{\kappa^{-2\alpha+1}}{2\alpha}.$$

Since $\{\psi_N^*\}_{N=1}^{\infty}$ is a strictly decreasing sequence with $\psi_1^* = 1 - \frac{\kappa^{-2\alpha+1}}{2\alpha}$. Therefore, for any fixed $d \geq 1$ and $N \geq 2$,

$$\lim_{n\to\infty} \frac{R_n^{\mathrm{MA}}(\mathbf{w}_{n,dN}^*)}{R_n^{\mathrm{MA}}(\mathbf{w}_{n,d}^*)} < 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{R_n^{\mathrm{MA}}(\mathbf{w}_{n,N}^*)}{R_n^{\mathrm{MS}}(m_n^*)} < 1.$$

□

Lemma S.3. *Continued to Lemma S.2, we have*

(iii) For two given $\gamma^*_{n,q_n} < z_1 < z_2 < 1$*, if there exist constants* $k > 1$*,* $\delta \geq \frac{z_1/(1-z_1)}{z_2/(1-z_2)}$*, and* $K > 1$ *such that* $\theta_{n,\lfloor kl_n \rfloor}/\theta_{n,l_n} \geq \delta$ *for any* $n \geq K$ *and any integer sequence* $\{l_n\}$ *satisfying* $\lim_{n\to\infty} l_n = \infty$*, then* $m_n(z_1) - m_n(z_2) \asymp m_n(z_1)$*.*

Proof. By using the condition of (iii) and the definition of $m_n(z)$, we have

$$\theta_{n,m_n(z_1)+1} \leq \frac{z_1}{(1-z_1)n} \leq \frac{z_2}{(1-z_2)n}\delta < \delta\theta_{n,m_n(z_2)} \leq \theta_{n,\lfloor km_n(z_2)\rfloor},$$

which yields that $m_n(z_1) \geq \lfloor km_n(z_2) \rfloor$. Thus, we have

$$m_n(z_1) - m_n(z_2) \geq \lfloor km_n(z_2) \rfloor - m_n(z_2) > (k-1)m_n(z_2) - 1.$$

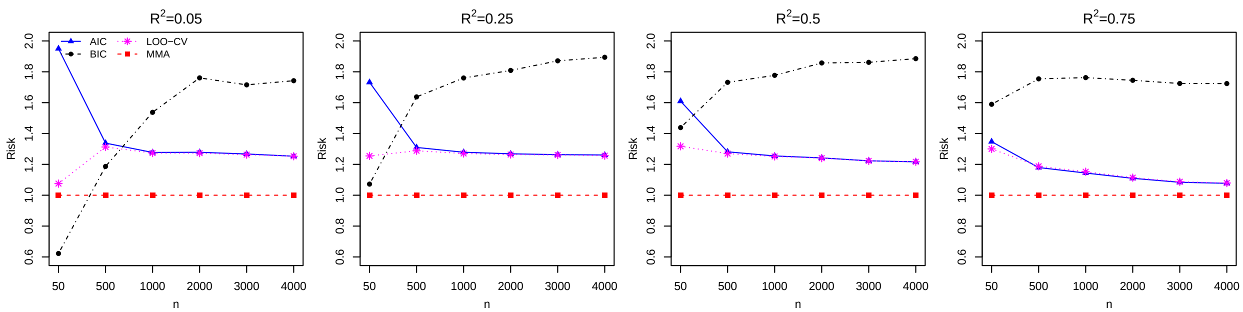
Therefore, $m_n(z_1) - m_n(z_2) \asymp m_n(z_1)$. We complete the proof of Lemma S.3. □

References

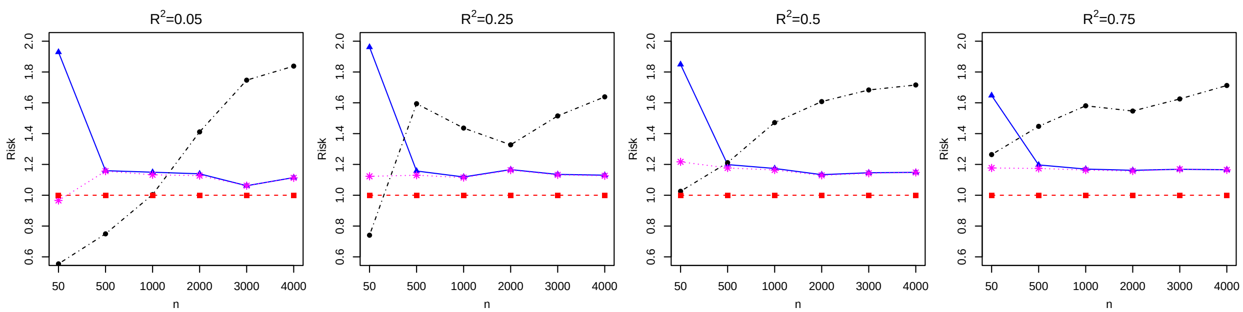
Bernstein, D. S. (2005). *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory*. Princeton University Press, Princeton.

Peng, J. and Yang, Y. (2022). On improvability of model selection by model averaging. *Journal of Econometrics*, 229(2):246–262.

S.4 Additional Figures in Section 6

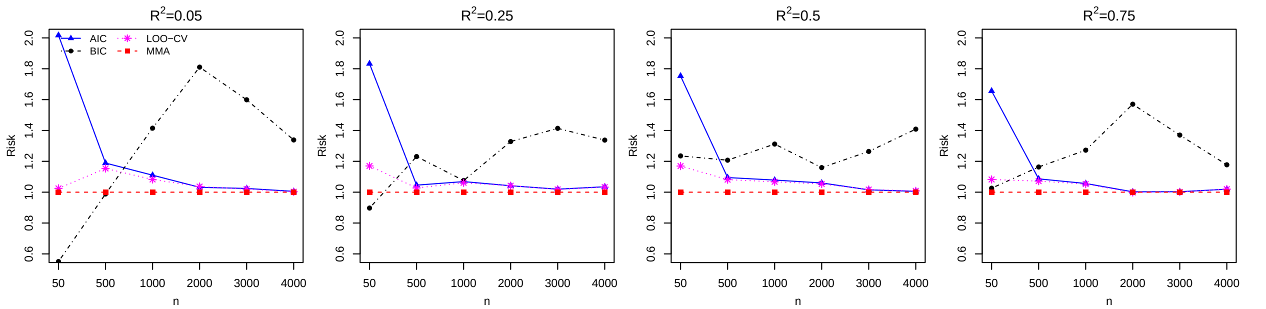


(a)

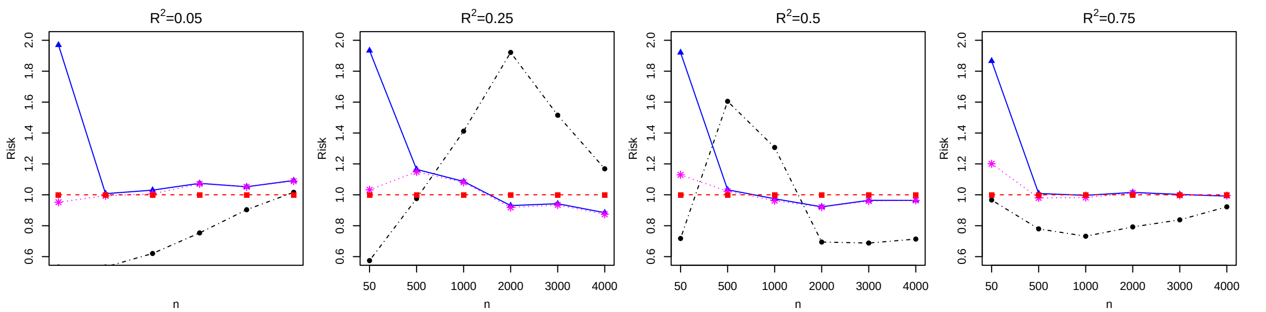


(b)

Figure S.2: Simulation results for Example 1 for the case of slowly decaying θ_m^*. Normalized risk functions for AIC, BIC, LOO-CV, and MMA when $\theta_m^* = m^{-2\alpha_1}/\sigma^2$ with $\alpha_1 = 1$ in row (a) and $\alpha_1 = 2$ in row (b).

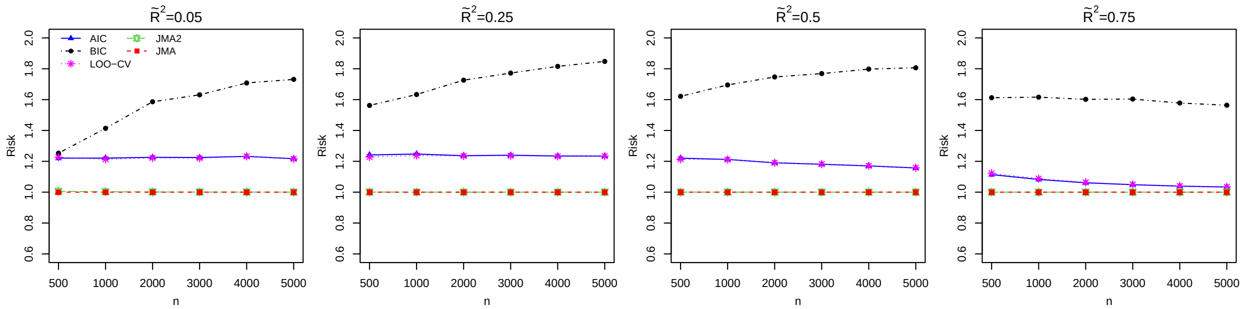


(a)

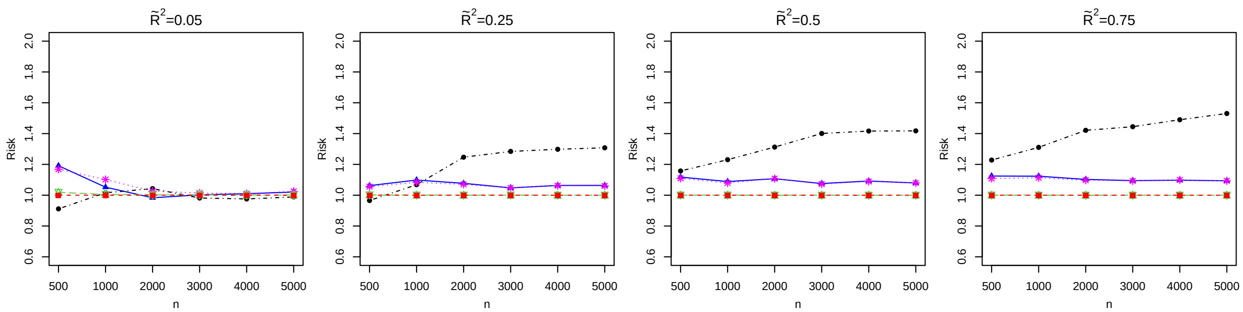


(b)

Figure S.3: Simulation results for Example 1 for the case of fast decaying θ_m^*. Normalized risk functions for AIC, BIC, LOO-CV, and MMA when $\theta_m^* = \exp(-2\alpha_2 m)/\sigma^2$ with $\alpha_2 = 1$ in row (a) and $\alpha_2 = 2$ in row (b).

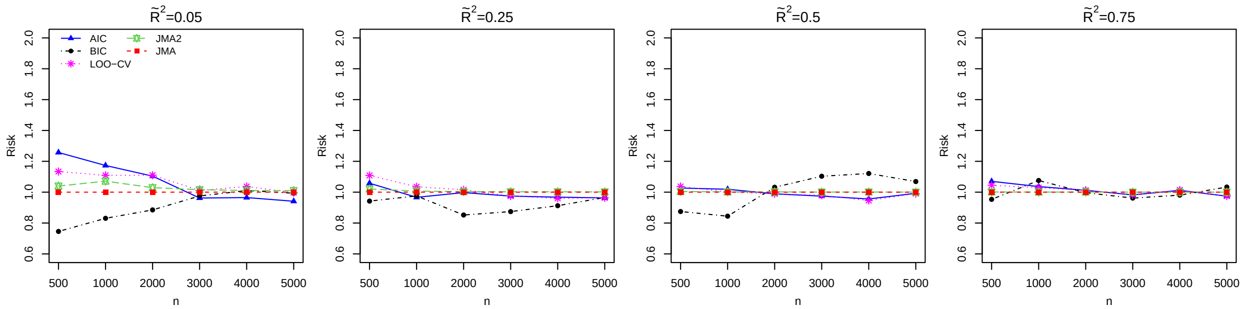


(a)

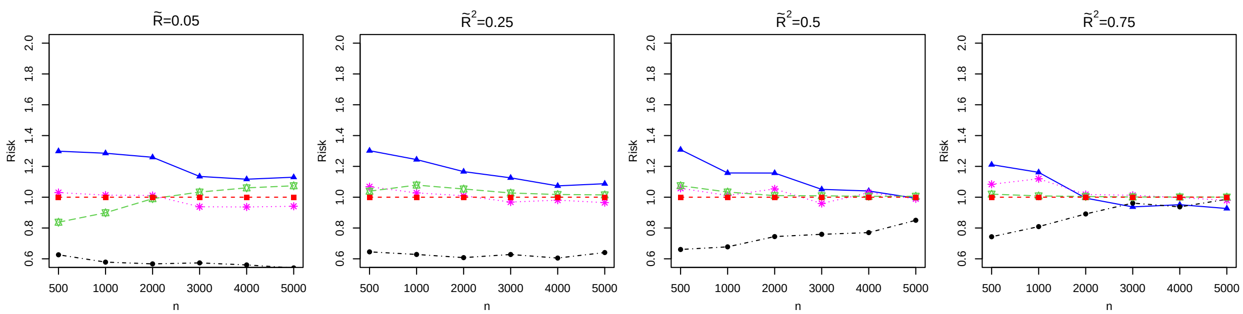


(b)

Figure S.4: Simulation results for Example 2 for the case of slowly decaying θ_m^*. Normalized risk functions for AIC, BIC, LOO-CV, JMA2, and JMA when $\theta_m^* = c^2 m^{-2\alpha_1}$ with $\alpha_1 = 1$ in row (a) and $\alpha_1 = 2$ in row (b).

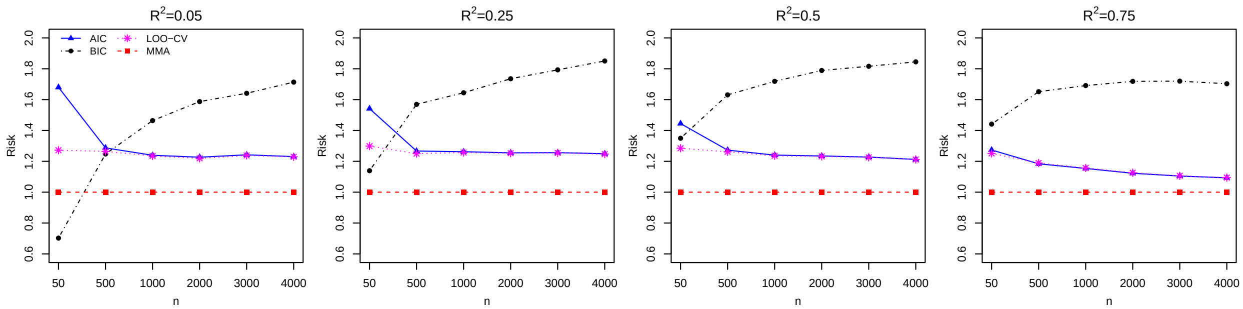


(a)

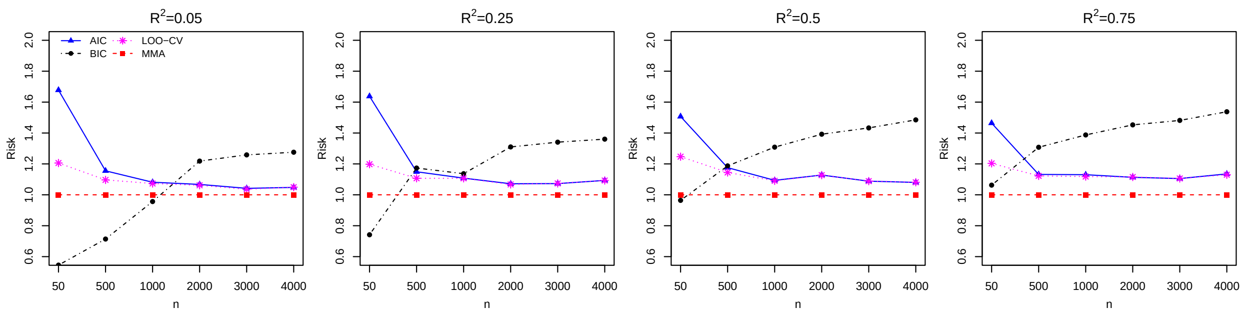


(b)

Figure S.5: Simulation results for Example 2 for the case of fast decaying θ_m^*. Normalized risk functions for AIC, BIC, LOO-CV, JMA2, and JMA when $\theta_m^* = c^2 \exp(-2\alpha_2 m)$ with $\alpha_2 = 1$ in row (a) and $\alpha_2 = 2$ in row (b).

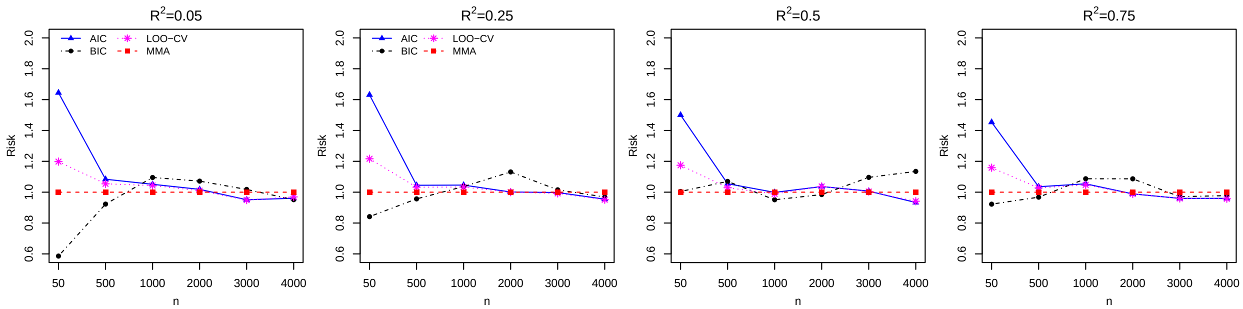


(a)

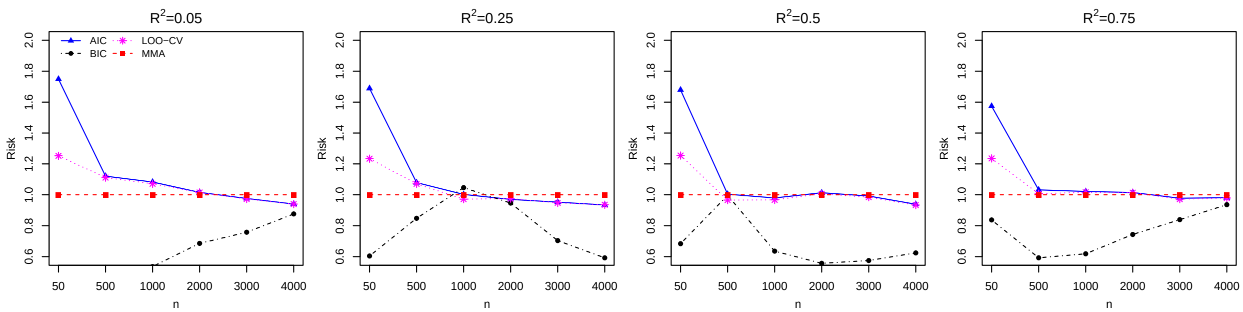


(b)

Figure S.6: Simulation results for Example 3 for the case of slowly decaying θ_m^*. Normalized risk functions for AIC, BIC, LOO-CV, and MMA when $\theta_m^* = m^{-2\alpha_1}/\sigma^2$ with $\alpha_1 = 1$ in row (a) and $\alpha_1 = 2$ in row (b).



(a)



(b)

Figure S.7: Simulation results for Example 3 for the case of fast decaying θ_m^*. Normalized risk functions for AIC, BIC, LOO-CV, and MMA when $\theta_m^* = \exp(-2\alpha_2 m)/\sigma^2$ with $\alpha_2 = 1$ in row (a) and $\alpha_2 = 2$ in row (b).

S.5 More Simulation Studies

We further design the following Example 4 to illustrate Corollary 1 under Condition M1.

Example 4 (Small number of candidate models) The setting of this example is the same as Peng and Yang (2022) except for the number of candidate models. We consider two cases with different decaying orders of $\theta_m^* = \beta_m^2/\sigma^2$:

- Case 1 (With θ_m^* satisfying Condition B1). Here, $\beta_m = m^{-\alpha_1}$, and α_1 is set to be 1, 1.5, or 2.

- Case 2 (With θ_m^* satisfying Condition B2). Here, $\beta_m = \exp(-\alpha_2 m)$, and α_2 is set to be 1, 1.5, or 2.

Note that $m_n^{**} \sim (\frac{n}{\sigma^2})^{\frac{1}{2\alpha_1}}$ in Case 1 and $m_n^{**} \sim \frac{1}{2\alpha_2} \log\left(\frac{n}{\sigma^2}\right)$ in Case 2. In order to illustrate Corollary 1 under Condition M1, M_n should be set to be small compared to m_n^{**}. Therefore, M_n is set to be $(\frac{n}{\sigma^2})^{\frac{1}{2\alpha_1}-\frac{1}{10}}$ in Case 1 and $\log\log\left(\frac{n}{\sigma^2}\right)$ in Case 2. It is easy to see that $\lim_{n\to\infty} M_n/m_n^{**} = 0$, thus Condition M1 holds for these two cases. For Case 1, the sample size n varies at 50, 500, 1000, 2000, 3000, and 4000. For Case 2, n varies at 50, 1000, 4000, 6000, 8000, and 10000.

Simulation results are summarized in Figures S.8 and S.9. In each figure, the simulation results with three coefficient decaying orders are displayed in rows (a), (b), and (c). In both the slowly decaying and fast decaying θ_m^* cases, the performance gap between AIC (or LOO-CV) and MMA becomes very close when n is large, which are consistent with the results of Corollary 1 under Condition M1.

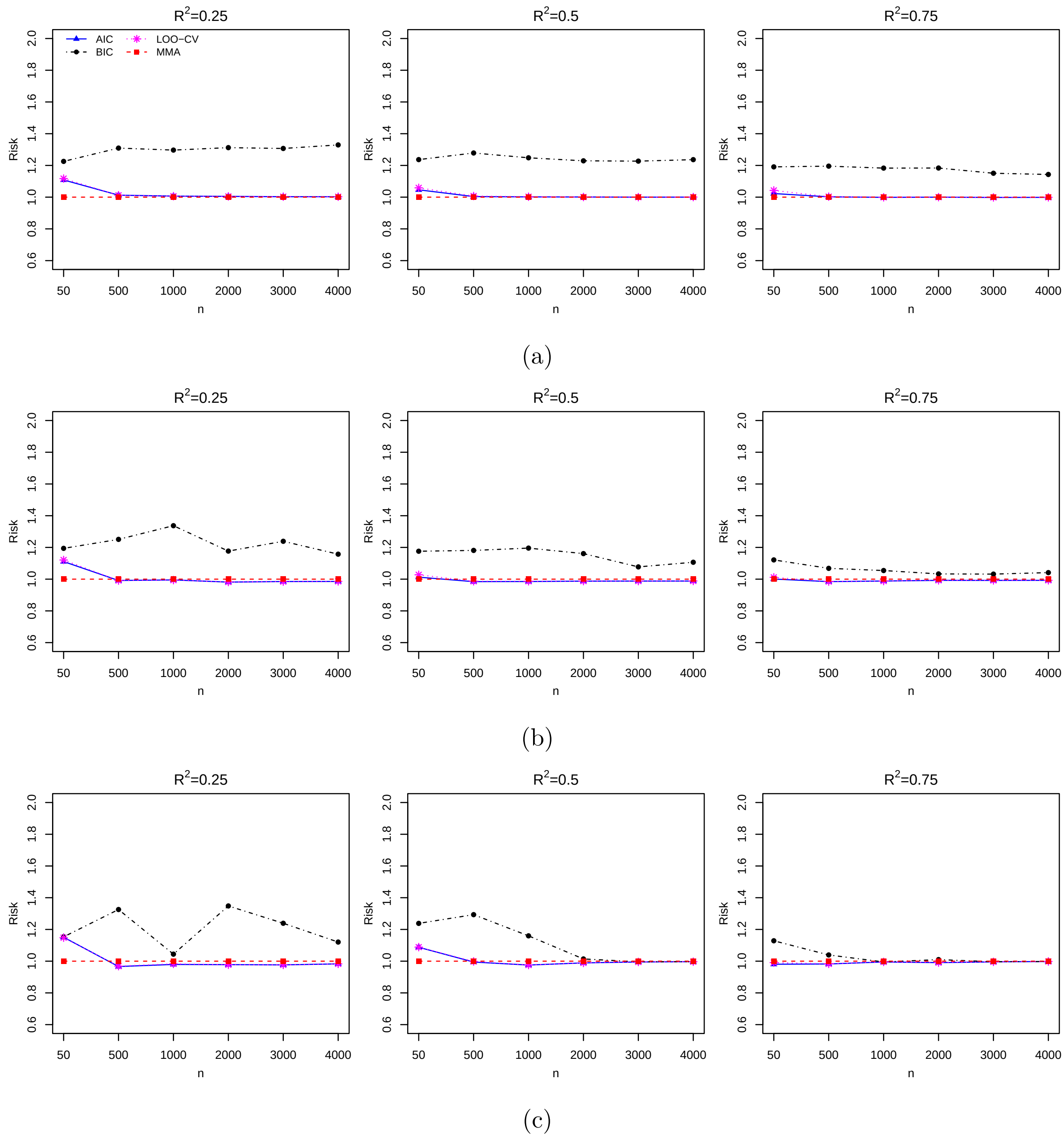
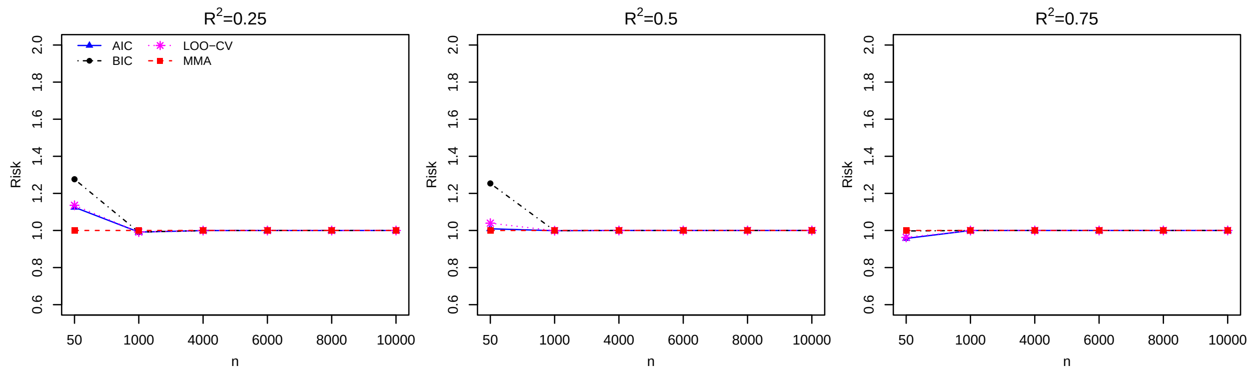
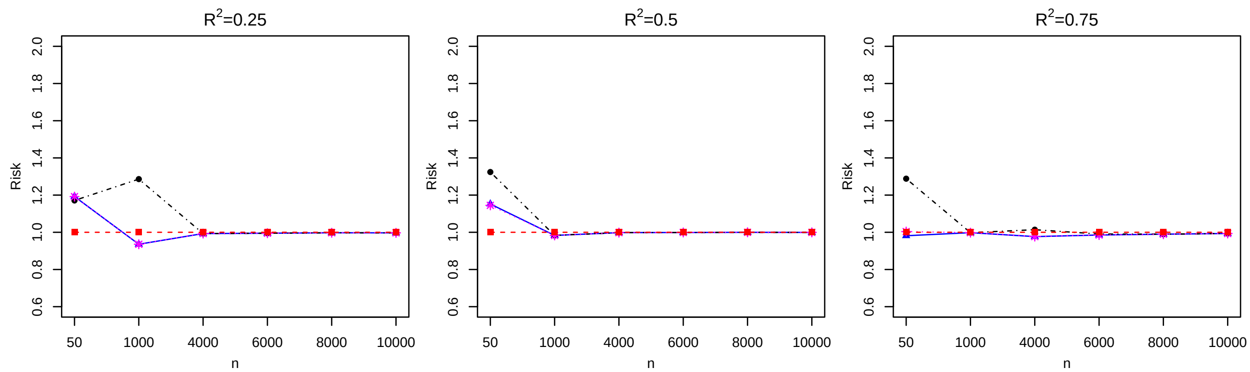


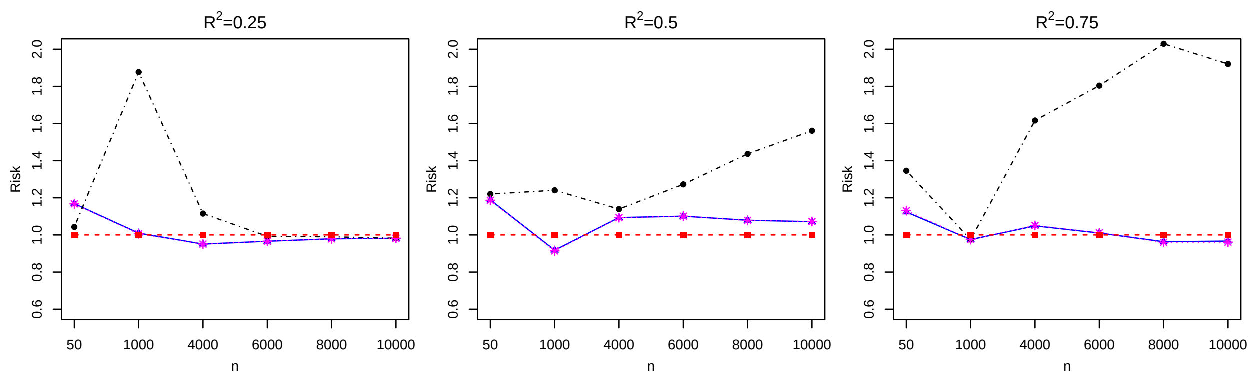
Figure S.8: Simulation results for Example 4 for the case of slowly decaying θ^*_m. Normalized risk functions for AIC, BIC, LOO-CV, and MMA when $\beta_m = m^{-\alpha_1}$ with $\alpha_1 = 1$ in row (a), $\alpha_1 = 1.5$ in row (b), and $\alpha_1 = 2$ in row (c).



(a)



(b)



(c)

Figure S.9: Simulation results for Example 4 for the case of fast decaying θ_m^*. Normalized risk functions for AIC, BIC, LOO-CV, and MMA when $\beta_m = \exp(-\alpha_2 m)$ with $\alpha_2 = 1$ in row (a), $\alpha_2 = 1.5$ in row (b), and $\alpha_2 = 2$ in row (c).