HOMEWORK SOLUTIONS #7

1. (a) For the case $\mathbf{Q} = 10^{-5}\mathbf{I}$, we find that $\lambda_{\text{opt}} = \texttt{0.9326}$ and $\lim_{n\to\infty} J(n)\big|_{\text{RLS}} = \texttt{0.002450}$ (or `-26.108 dB`), and that $\mu_{\text{opt}} = \texttt{0.0742}$ and $\lim_{n\to\infty} J(n)\big|_{\text{LMS}} = \texttt{0.002450}$ (or `-26.108 dB`) which is identical to the RLS case.

   (b) For the case $\mathbf{Q} = 10^{-5}\mathbf{I}$, Fig. 1 and Fig. 2 show parameter trajectories and learning curves, respectively. Averaging the experimental values, we find $\lim_{n\to\infty} J(n)\big|_{\text{RLS}} = \texttt{0.002496}$ (or `-26.027 dB`), and that $\lim_{n\to\infty} J(n)\big|_{\text{LMS}} = \texttt{0.002504}$ (or `-26.014 dB`), which shows that the experimental values match the theoretical quite well. As expected, LMS and RLS have almost exactly the same steady-state performance. As can be seen from the plots, however, RLS converges much faster than LMS.

   (c) For the case $\mathbf{Q} = 10^{-5}\mathbf{R}$, we find that $\lambda_{\text{opt}} = \texttt{0.9147}$ and $\lim_{n\to\infty} J(n)\big|_{\text{RLS}} = \texttt{0.002587}$ (or `-25.873 dB`), and that $\mu_{\text{opt}} = \texttt{0.0707}$ and $\lim_{n\to\infty} J(n)\big|_{\text{LMS}} = \texttt{0.002427}$ (or `-26.150 dB`) which is slightly lower than for RLS. Fig. 1 and Fig. 2 show parameter trajectories and learning curves, respectively. Averaging the experimental values, we find $\lim_{n\to\infty} J(n)\big|_{\text{RLS}} = \texttt{0.002611}$ (or `-25.832 dB`), and that $\lim_{n\to\infty} J(n)\big|_{\text{LMS}} = \texttt{0.002495}$ (or `-26.029 dB`), which shows that the experimental values match the theoretical quite well. As expected, LMS has slightly superior steady-state performance relative to RLS. As can be seen from the plots, however, RLS converges much faster than LMS.

   (d) For the case $\mathbf{Q} = 10^{-5}\mathbf{R}^{-1}$, we find that $\lambda_{\text{opt}} = \texttt{0.9293}$ and $\lim_{n\to\infty} J(n)\big|_{\text{RLS}} = \texttt{0.002475}$ (or `-26.065 dB`), and that $\mu_{\text{opt}} = \texttt{0.1108}$ and $\lim_{n\to\infty} J(n)\big|_{\text{LMS}} = \texttt{0.002712}$ (or `-25.667 dB`) which is slightly higher than for RLS. Fig. 1 and Fig. 2 show parameter trajectories and learning curves, respectively. Averaging the experimental values, we find $\lim_{n\to\infty} J(n)\big|_{\text{RLS}} = \texttt{0.002490}$ (or `-26.037 dB`), and that $\lim_{n\to\infty} J(n)\big|_{\text{LMS}} = \texttt{0.002686}$ (or `-25.709 dB`), which shows that the experimental values match the theoretical quite well. As expected, LMS has slightly inferior steady-state performance relative to RLS. Again, as can be seen from the plots RLS converges much faster than LMS.

2. (a) Here we need to prove that $\text{tr}(\mathbf{R})^2 \le M\,\text{tr}(\mathbf{R}^2)$. The eigendecomposition $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ implies $\mathbf{R}^2 = \mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^H$, so that $\text{tr}(\mathbf{R}) = \sum_i \lambda_i$ and $\text{tr}(\mathbf{R}^2) = \sum_i \lambda_i^2$, where λ_i are the eigenvalues of $\mathbf{R}$. Now define

$$\begin{aligned}\boldsymbol{\lambda} &= [\lambda_1, \lambda_2, \ldots, \lambda_M]^t \\ \mathbf{1} &= [1, 1, \ldots, 1]^t\end{aligned}$$

   noting that $\sum_i \lambda_i = \mathbf{1}^t\boldsymbol{\lambda}$, $\|\mathbf{1}\|^2 = M$, and $\|\boldsymbol{\lambda}\|^2 = \sum_i \lambda_i^2$. Then from Cauchy-Schwarz and the fact that $\lambda_i \ge 0$,

$$\begin{aligned} |\mathbf{1}^t\boldsymbol{\lambda}|^2 &\le \|\mathbf{1}\|^2\|\boldsymbol{\lambda}\|^2 \\ \Rightarrow \left(\sum_i \lambda_i\right)^2 &\le M\sum_i \lambda_i^2 \\ \Rightarrow \text{tr}(\mathbf{R})^2 &\le M\,\text{tr}(\mathbf{R}^2)\end{aligned}$$

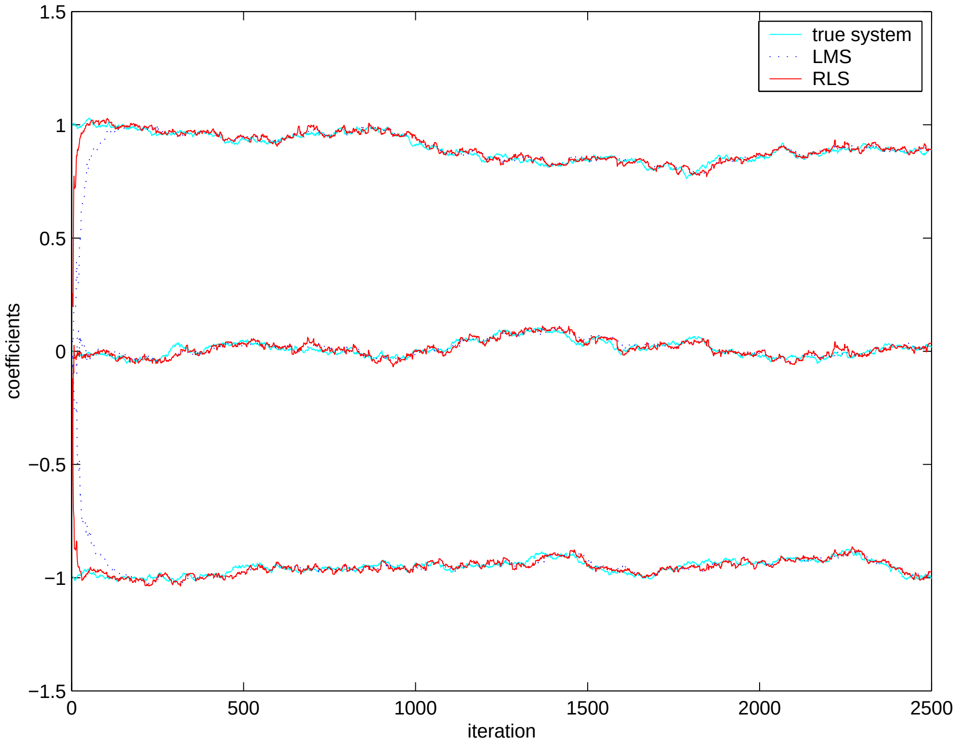

Figure 1: Parameter trajectories, $\mathbf{Q} = 10^{-5}\mathbf{I}$.

(b) Here we need to prove that $\operatorname{tr}(\mathbf{R})\operatorname{tr}(\mathbf{R}^{-1}) \geq M^2$. Since $\lambda_i > 0$ we can define

$$\begin{aligned} \mathbf{x} &= \left[\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_M}\right]^t \\ \mathbf{y} &= \left[\left(\sqrt{\lambda_1}\right)^{-1}, \left(\sqrt{\lambda_2}\right)^{-1}, \ldots, \left(\sqrt{\lambda_M}\right)^{-1}\right]^t \end{aligned}$$

Then we see that $\|\mathbf{x}\|^2 = \sum_i \lambda_i = \operatorname{tr}(\mathbf{R})$, that $\|\mathbf{y}\|^2 = \sum_i \lambda_i^{-1} = \operatorname{tr}(\mathbf{R}^{-1})$, and that $\mathbf{x}^t\mathbf{y} = M$. Then from Cauchy-Schwarz

$$\begin{aligned} |\mathbf{x}^t\mathbf{y}|^2 &\leq \|\mathbf{x}\|^2\|\mathbf{y}\|^2 \\ \Rightarrow \quad M^2 &\leq \operatorname{tr}(\mathbf{R})\operatorname{tr}(\mathbf{R}^{-1}) \end{aligned}$$

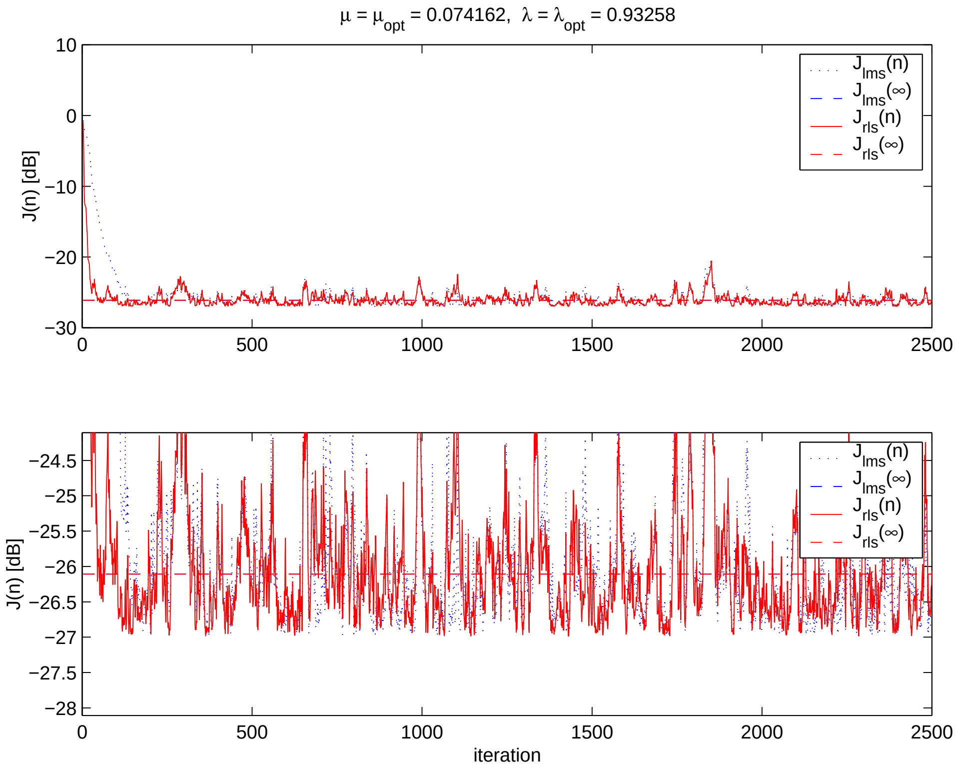

Figure 2: Learning curves, $\mathbf{Q} = 10^{-5}\mathbf{I}$.

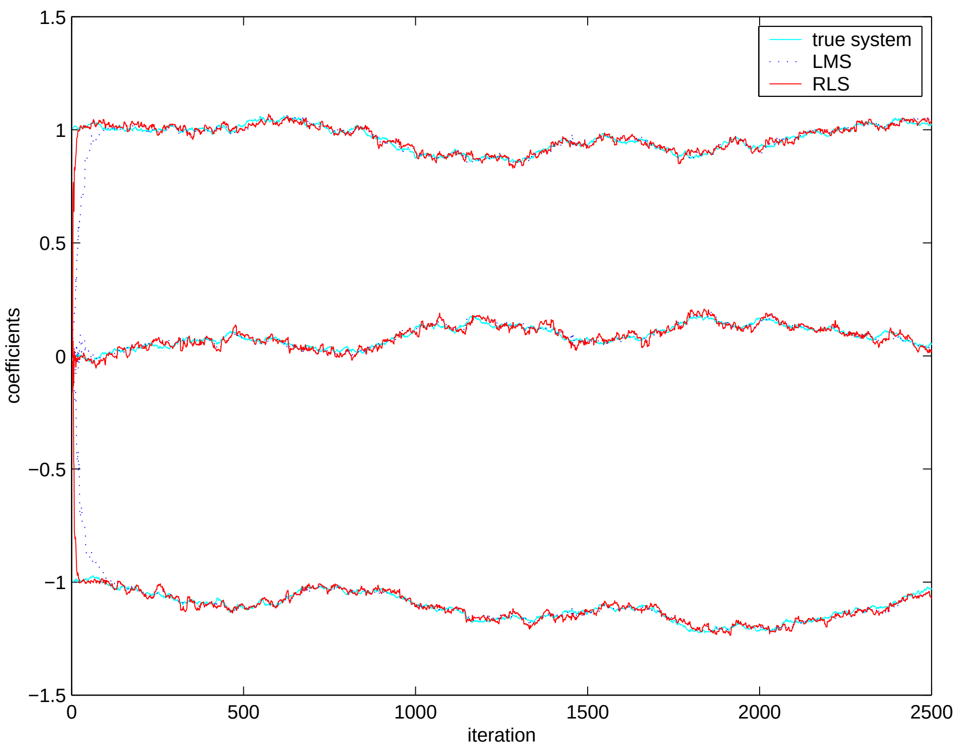

Figure 3: Parameter trajectories, $\mathbf{Q} = 10^{-5}\mathbf{R}$.

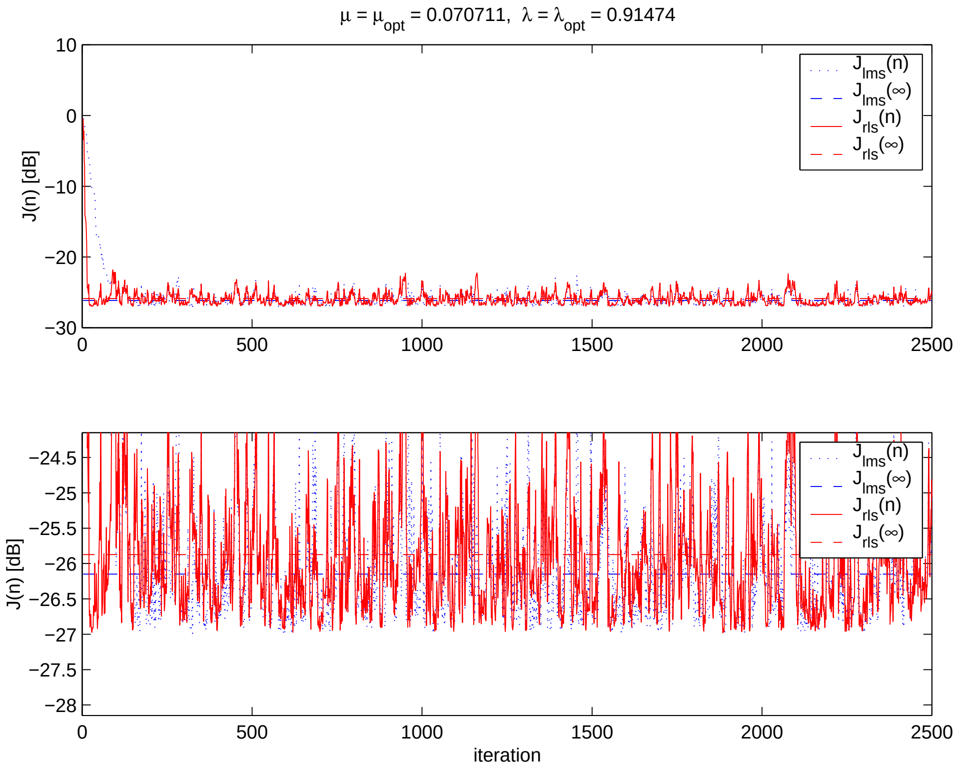

Figure 4: Learning curves, $\mathbf{Q} = 10^{-5}\mathbf{R}$.

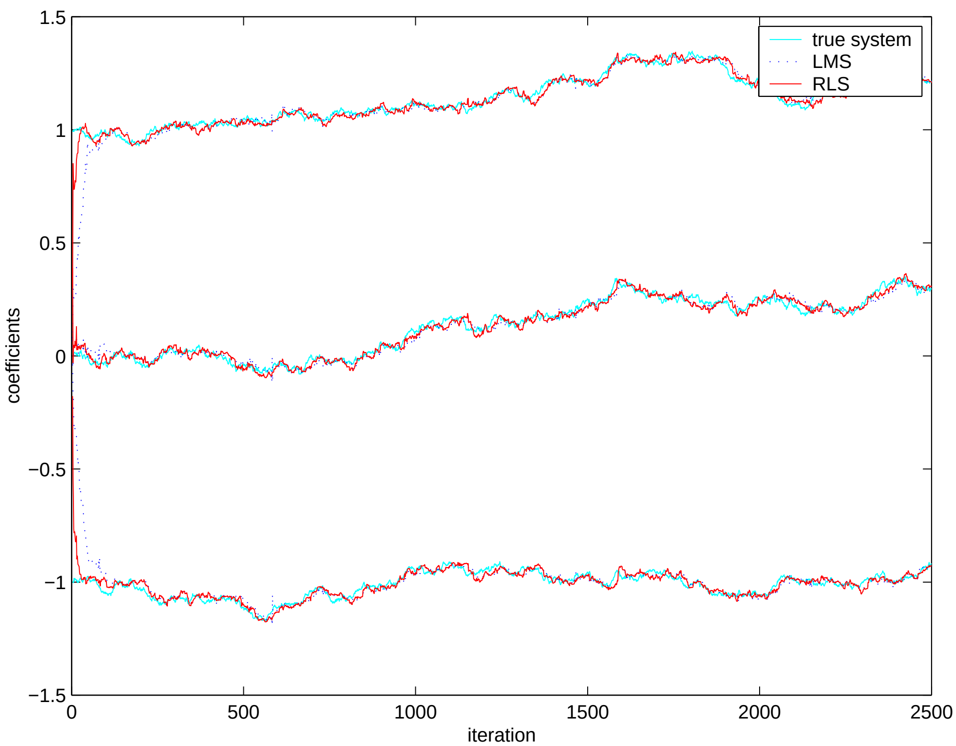

Figure 5: Parameter trajectories, $\mathbf{Q} = 10^{-5}\mathbf{R}^{-1}$.

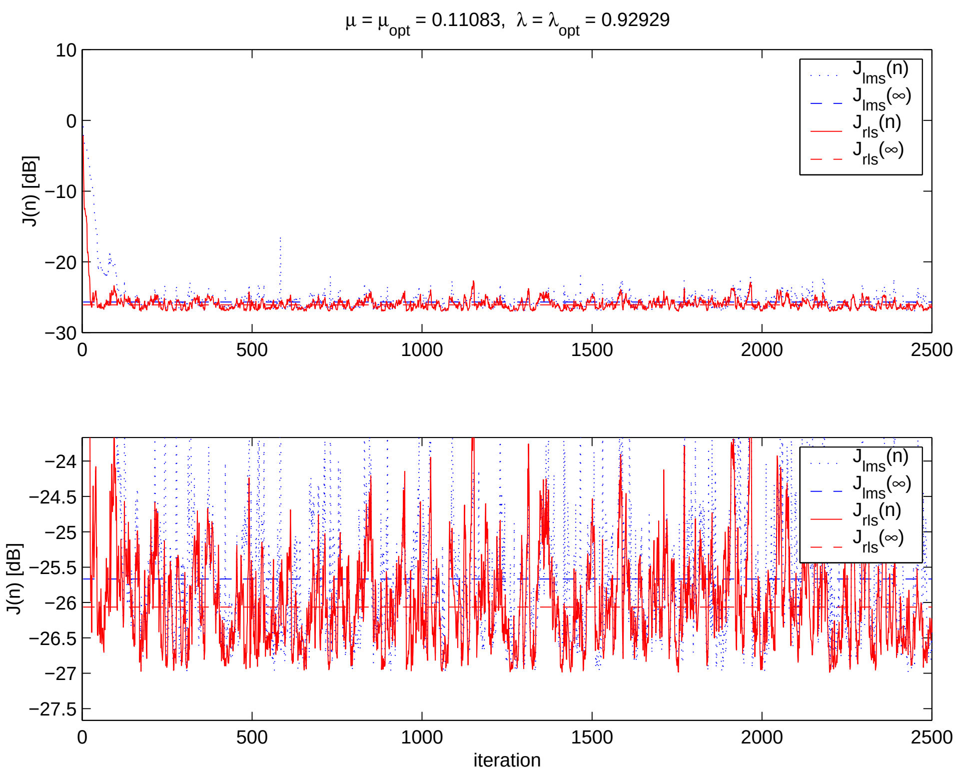

Figure 6: Learning curves, $\mathbf{Q} = 10^{-5}\mathbf{R}^{-1}$.