Supplementary Material: Global Convergence of Policy Gradient for Linear-Quadratic Mean-Field Control/Game in Continuous Time

Weichen Wang *, Jiequn Han †, Zhuoran Yang ‡ and Zhaoran Wang §

Abstract

The supplemental material contains supporting proofs for the main document.

A Proofs for Section 2

Lemma A.1. (Solution of continuous Lyapunov equation). Suppose W is stable. The solution Y of continuous Lyapunov equation

$$WY + YW^\top + Q = 0$$

can be written as

$$Y = \int_0^\infty e^{W\tau} Q e^{W^\top \tau} \mathrm{d}\tau. \tag{A.1}$$

Proof. The result can be found in Theorem 7.5 of [1], so we omit its proof. □

In the following, given K such that $A - BK$ is stable, we define two operators $\mathcal{T}_K, \mathcal{F}_K$ on symmetric matrix X as

$$\mathcal{T}_K(X) \coloneqq \int_0^\infty e^{(A-BK)\tau} X e^{(A-BK)^\top \tau} \mathrm{d}\tau,$$
$$\mathcal{F}_K(X) \coloneqq (A - BK)X + X(A - BK)^\top.$$

Then

$$\mathcal{F}_K \circ \mathcal{T}_K + I = 0,$$

or

$$\mathcal{T}_K = -\mathcal{F}_K^{-1}.$$

Additionally, from (4) we have

$$\Sigma_K = \mathcal{T}_K(DD^\top).$$

*Faculty of Business and Economics, The University of Hong Kong. <nickweichwang@gmail.com>
†Department of Mathematics, Princeton University. <jiequnh@princeton.edu>
‡Department of Operations Research & Financial Engineering, Princeton University. <zy6@princeton.edu>
§Department of Industrial Engineering & Management Sciences, Northwestern University. <zhaoran.wang@northwestern.edu>

Lemma A.2. (Perturbation of P_K). Assume K, K' are both stable. Then

$$P_{K'} - P_K = \int_0^\infty e^{(A-BK')^\top \tau}[E_K^\top (K'-K) + (K'-K)^\top E_K + (K'-K)^\top R(K'-K)]e^{(A-BK')\tau}\mathrm{d}\tau. \quad \text{(A.2)}$$

Moreover, this implies that P_K is differentiable.

Proof. Taking the difference between two equations (7) corresponding to K' and K, we have

$$\begin{aligned}
0 &= (A-BK')^\top P_{K'} + P_{K'}(A-BK')^\top - (A-BK'+B(K-K'))^\top P_K + P_K(A-BK'+B(K-K'))^\top \\
&\quad + (K'-K+K)^\top R(K'-K+K) - K^\top RK \\
&= (A-BK')^\top (P_{K'}-P_K) + (P_{K'}-P_K)(A-BK')^\top - (K'-K)^\top B^\top P_K - P_K B(K'-K) \\
&\quad + (K'-K+K)^\top R(K'-K+K) - K^\top RK \\
&= (A-BK')^\top (P_{K'}-P_K) + (P_{K'}-P_K)(A-BK')^\top \\
&\quad + E_K^\top (K'-K) + (K'-K)^\top E_K + (K'-K)^\top R(K'-K).
\end{aligned}$$

Here $E_K = RK - B^\top P_K$ is defined in Proposition 1. In other words, $P_{K'} - P_K$ is the solution of the continuous Lyapunov equation

$$(A-BK')^\top Y + Y(A-BK') + E_K^\top (K'-K) + (K'-K)^\top E_K + (K'-K)^\top R(K'-K) = 0,$$

in which Y is the unknown matrix. Recalling Lemma A.1, we finish the first part of the proof.

Define vectorization operator for $n \times m$ matrix $Y = (y_{ij})_{i \le n, j \le m}$ as

$$\mathrm{vec}(Y) = (y_{11}, \ldots, y_{n1}, y_{12}, \ldots, y_{n2}, \ldots, y_{1m}, \ldots, y_{nm})^\top .$$

We have the fact that $\mathrm{vec}(ABC) = (C^\top \otimes A)\,\mathrm{vec}(B)$. Using this, (A.2) gives us

$$\begin{aligned}
\mathrm{vec}(P_{K'} - P_K) &= \int_0^\infty \mathrm{vec}\left(e^{(A-BK')^\top \tau}[E_K^\top (K'-K) + (K'-K)^\top E_K + (K'-K)R(K'-K)]e^{(A-BK')\tau}\right)\mathrm{d}\tau \\
&= \int_0^\infty \left(e^{(A-BK')^\top \tau} \otimes e^{(A-BK')^\top \tau}\mathrm{d}\tau\right) \mathrm{vec}[E_K^\top (K'-K) + (K'-K)^\top E_K + (K'-K)R(K'-K)] \\
&= \int_0^\infty \left(e^{(A-BK')^\top \tau} \otimes e^{(A-BK')^\top \tau}\mathrm{d}\tau\right) \mathrm{vec}[E_{K'}^\top (K'-K) + (K'-K)^\top E_{K'} + U]\,,
\end{aligned}$$

where

$$\begin{aligned}
U &= (K'-K)^\top R(K'-K) + (E_K - E_{K'})^\top (K'-K) + (K'-K)^\top (E_K - E_{K'}) \\
&= -(K'-K)^\top R(K'-K) + (P_{K'} - P_K)B(K'-K) + (K'-K)^\top B^\top (P_{K'} - P_K) \\
&= O(\|K'-K\|_F^2)\,.
\end{aligned}$$

The last line uses the expression of $P_{K'} - P_K$ in the first part of Lemma A.2 again. Therefore, there exists $Z_{K'}$ that depend on $A - BK'$ and $E_{K'}$ such that $\mathrm{vec}(P_K - P_{K'}) = Z_{K'}\,\mathrm{vec}(K - K') + O(\|K - K'\|_F^2)$, where $Z_{K'}$ will be defined as the derivative of $\mathrm{vec}(P_K)$ at $K = K'$ with respect to $\mathrm{vec}(K)$. Therefore, P_K is indeed differentiable and its differential $\mathrm{d}P_K$ used in the proof of Proposition 1 below is well-defined. □

Now we are ready to prove the expression of the policy gradient as follows.

Proposition A.3. (Proposition 1).

$$\nabla_K J(K) = 2(RK - B^\top P_K)\Sigma_K = 2E_K \Sigma_K, \quad \text{(A.3)}$$

where $E_K = RK - B^\top P_K$.

Proof. Rewrite the Lyapunov equation (7) as $\phi(K, P_K) = 0$, where ϕ is a function of two independent arguments, defined as

$$\phi(K, P_K) := (A - BK)^\top P_K + P_K(A - BK) + Q + K^\top RK.$$

Taking differential on both sides (the differentiability of P_K has been shown in Lemma A.2), we have

$$\begin{aligned} 0 &= \nabla_K \phi(K, P_K)\mathrm{d}K + \nabla_{P_K} \phi(K, P_K)\mathrm{d}P_K \\ &= [(-B\mathrm{d}K)^\top P_K + P_K(-B\mathrm{d}K) + (\mathrm{d}K)^\top RK + K^\top R\mathrm{d}K] + [(A - BK)^\top \mathrm{d}P_K + \mathrm{d}P_K(A - BK)], \end{aligned}$$

or equivalently,

$$(A - BK)^\top \mathrm{d}P_K + \mathrm{d}P_K(A - BK) + (K^\top R - P_K B)\mathrm{d}K + (dK)^\top (RK - B^\top P_K) = 0. \tag{A.4}$$

Note that (4)(A.4) have similar structures. We apply the trace operator to (4) left multiplied by $\mathrm{d}P_K$ and (A.4) left multiplied by Σ_K, and then take the difference to obtain

$$\begin{aligned} \mathrm{tr}(\mathrm{d}P_K DD^\top) &= \mathrm{tr}[\Sigma_K(K^\top R - P_K B)\mathrm{d}K + \Sigma_K(dK)^\top (RK - B^\top P_K)] \\ &= \mathrm{tr}[2\Sigma_K(K^\top R - P_K B)\mathrm{d}K]. \end{aligned}$$

From (8), by definition, we have

$$\mathrm{tr}[(\nabla_K J(K))^\top \mathrm{d}K] = \mathrm{d}J(K) = \mathrm{tr}(\mathrm{d}P_K DD^\top).$$

Comparing the above two equations, since the matrix quantities are equal for any direction of $\mathrm{d}K$, we conclude $\nabla_K J(K) = 2(RK - B^\top P_K)\Sigma_K$. □

Lemma A.4. (Lemma 2). The cost function is gradient dominated [3], that is

$$J(K) - J(K^*) \leq \frac{\|\Sigma_{K^*}\|}{\sigma_{\min}(R)\sigma^2_{\min}(DD^\top)} \mathrm{tr}(\nabla_K J(K)^\top \nabla_K J(K)). \tag{A.5}$$

In additional, we have the following lower bound for $J(K) - J(K^*)$

$$J(K) - J(K^*) \geq \frac{\sigma_{\min}(DD^\top)}{\|R\|} \mathrm{tr}(E_K^\top E_K). \tag{A.6}$$

Proof. Based on (8) and Lemma A.2, we have

$$\begin{aligned} & J(K') - J(K) \\ =\ & \mathrm{tr}[(P_{K'} - P_K)DD^\top] \\ =\ & \mathrm{tr}\left[\int_0^\infty e^{(A-BK')^\top \tau}[E_K^\top (K' - K) + (K' - K)^\top E_K + (K' - K)^\top R(K' - K)]e^{(A-BK')\tau} DD^\top \mathrm{d}\tau\right] \\ =\ & \mathrm{tr}\left[\int_0^\infty e^{(A-BK')\tau} DD^\top e^{(A-BK')^\top \tau}\mathrm{d}\tau[E_K^\top (K' - K) + (K' - K)^\top E_K + (K' - K)^\top R(K' - K)]\right] \\ =\ & \mathrm{tr}[\Sigma_{K'}[E_K^\top (K' - K) + (K' - K)^\top E_K + (K' - K)^\top R(K' - K)]] \\ =\ & \mathrm{tr}[\Sigma_{K'}[(K' - K + R^{-1}E_K)^\top R(K' - K + R^{-1}E_K) - E_K^\top R^{-1} E_K]]. \end{aligned}$$

Here the second equality uses Lemma A.2; the fourth equality uses the fact that $\Sigma_{K'}$ is the solution the Lyapunov equation $(A - BK')X + X(A - BK)^\top + DD^\top = 0$ and Lemma A.1.

To prove the upper bound (A.5), we use the fact that the quadratic term $(K' - K + R^{-1}E_K)^\top R(K' - K + R^{-1}E_K)$ above is positive semi-definite. Letting $K' = K^*$, we have

$$\begin{aligned} J(K) - J(K^*) &= \operatorname{tr}[\Sigma_{K^*}[E_K^\top R^{-1}E_K - (K^* - K + R^{-1}E_K)^\top R(K^* - K + R^{-1}E_K)]] \\ &\leq \operatorname{tr}[\Sigma_{K^*}E_K^\top R^{-1}E_K] \\ &\leq \frac{\|\Sigma_{K^*}\|}{\sigma_{\min}(R)} \operatorname{tr}(E_K^\top E_K) \\ &\leq \frac{\|\Sigma_{K^*}\|}{\sigma_{\min}(R)\sigma_{\min}^2(\Sigma_K)} \operatorname{tr}(\nabla_K J(K)^\top \nabla_K J(K)) \\ &\leq \frac{\|\Sigma_{K^*}\|}{\sigma_{\min}(R)\sigma_{\min}^2(DD^\top)} \operatorname{tr}(\nabla_K J(K)^\top \nabla_K J(K)). \end{aligned}$$

The last inequality follows from the fact that $\Sigma_K \succeq DD^\top \succeq \sigma_{\min}(DD^\top) \cdot I_d$.

To prove the lower bound, we choose a specific form of K' to make the quadratic term to be zero and use the fact that $J(K^*) \leq J(K')$. Letting $K' = K - R^{-1}E_K$, we have

$$J(K) - J(K') = \operatorname{tr}[\Sigma_{K'}E_K^\top R^{-1}E_K].$$

Then

$$\begin{aligned} J(K) - J(K^*) &\geq J(K) - J(K') \\ &\geq \operatorname{tr}[\Sigma_{K'}E_K^\top R^{-1}E_K] \\ &\geq \frac{\sigma_{\min}(DD^\top)}{\|R\|} \operatorname{tr}(E_K^\top E_K). \end{aligned}$$

□

Lemma A.5. (Perturbation analysis of Σ_K) Suppose $A - BK$ is stable and

$$\|K' - K\| \leq \frac{\sigma_{\min}(Q)\sigma_{\min}(DD^\top)}{4J(K)\|B\|},$$

then $A - BK'$ is also stable and

$$\|\Sigma_{K'} - \Sigma_K\| \leq 4\left(\frac{J(K)}{\sigma_{\min}(Q)}\right)^2 \frac{\|B\|}{\sigma_{\min}(DD^\top)}\|K' - K\|.$$

Proof. The first claim is easy to prove with Lemma 10 in [4]. The second claim is similar to Appendix C.4 in [2]. We first claim

$$\|\Sigma_K\| \leq \frac{J(K)}{\sigma_{\min}(Q)} \text{ and } \|\mathcal{T}_K\| \leq \frac{\|\Sigma_K\|}{\sigma_{\min}(DD^\top)}, \tag{A.7}$$

and it is clear to see that

$$\|\mathcal{F}_{K'} - \mathcal{F}_K\| \leq 2\|B\|\|K' - K\|.$$

Then

$$\|\mathcal{T}_K\|\|\mathcal{F}_{K'} - \mathcal{F}_K\| \leq \frac{2J(K)\|B\|\|K' - K\|}{\sigma_{\min}(Q)\sigma_{\min}(DD^\top)} \leq \frac{1}{2}.$$

Then we have

$$\begin{aligned}\|\Sigma_{K'} - \Sigma_K\| &= \|(\mathcal{T}_{K'} - \mathcal{T}_K)(DD^\top)\| \leq \|\mathcal{T}_K\| \|\mathcal{F}_{K'} - \mathcal{F}_K\| \|\Sigma_{K'}\| \\ &\leq \|\mathcal{T}_K\| \|\mathcal{F}_{K'} - \mathcal{F}_K\| (\|\Sigma_K\| + \|\Sigma_{K'} - \Sigma_K\|)\end{aligned}$$

Therefore,

$$\begin{aligned}\|\Sigma_{K'} - \Sigma_K\| &\leq 2\|\mathcal{T}_K\| \|\mathcal{F}_{K'} - \mathcal{F}_K\| \|\Sigma_K\| \\ &\leq 4 \left(\frac{J(K)}{\sigma_{\min}(Q)} \right)^2 \frac{\|B\|}{\sigma_{\min}(DD^\top)} \|K' - K\|.\end{aligned}$$

So it remains to show the claim in (A.7). The first claim can be seen from

$$J(K) = \operatorname{tr}(\Sigma_K(Q + K^\top RK)) \geq \operatorname{tr}(\Sigma_K)\sigma_{\min}(Q) \geq \|\Sigma_K\| \sigma_{\min}(Q)\,.$$

The second claim can be shown from the following fact. For any unit vector $v \in \mathbb{R}^d$ and unit spectral norm matrix X,

$$\begin{aligned}v^\top \mathcal{T}_K(X) v &= \int_0^\infty \operatorname{tr}(X e^{(A-BK)^\top \tau} v v^\top e^{(A-BK)\tau}) \mathrm{d}\tau \\ &\leq \int_0^\infty \operatorname{tr}(DD^\top e^{(A-BK)^\top \tau} v v^\top e^{(A-BK)\tau}) \mathrm{d}\tau \cdot \|(DD^\top)^{-1/2} X (DD^\top)^{-1/2}\| \\ &= (v^\top \Sigma_K v) \cdot \|(DD^\top)^{-1/2} X (DD^\top)^{-1/2}\| \leq \|\Sigma_K\| \sigma_{\min}^{-1}(DD^\top)\,.\end{aligned}$$

We now complete the proof. □

Lemma A.6. (Estimate of one-step GD). Suppose $K' = K - \eta \nabla_K J(K)$ with

$$\eta \leq \min \left\{ \frac{3\sigma_{\min}(Q)}{8J(K)\|R\|}, \frac{1}{16} \left(\frac{\sigma_{\min}(Q)\sigma_{\min}(DD^\top)}{J(K)} \right)^2 \frac{1}{\|B\| \|\nabla_K J(K)\|} \right\},$$

then

$$J(K') - J(K^*) \leq \left(1 - \eta \frac{\sigma_{\min}(R)\sigma_{\min}^2(DD^\top)}{\|\Sigma_{K^*}\|} \right) (J(K) - J(K^*)).$$

Proof. By the proof of Lemma 2, we have

$$\begin{aligned}&J(K) - J(K') \\ &= 2\operatorname{tr}[\Sigma_{K'}(K - K')^\top E_K] - \operatorname{tr}[\Sigma_{K'}(K - K')^\top R(K - K')] \\ &= 4\eta \operatorname{tr}(\Sigma_{K'}\Sigma_K E_K^\top E_K) - 4\eta^2 \operatorname{tr}(\Sigma_K \Sigma_{K'} \Sigma_K E_K^\top R E_K) \\ &\geq 4\eta \operatorname{tr}(\Sigma_K E_K^\top E_K \Sigma_K) - 4\eta \|\Sigma_{K'} - \Sigma_K\| \operatorname{tr}(\Sigma_K E_K^\top E_K) - 4\eta^2 \|\Sigma_{K'}\| \|R\| \operatorname{tr}(\Sigma_K E_K^\top E_K \Sigma_K) \\ &\geq 4\eta \operatorname{tr}(\Sigma_K E_K^\top E_K \Sigma_K) - 4\eta \frac{\|\Sigma_{K'} - \Sigma_K\|}{\sigma_{\min}(\Sigma_K)} \operatorname{tr}(\Sigma_K E_K^\top E_K \Sigma_K) - 4\eta^2 \|\Sigma_{K'}\| \|R\| \operatorname{tr}(\Sigma_K E_K^\top E_K \Sigma_K) \\ &= 4\eta \left(1 - \frac{\|\Sigma_{K'} - \Sigma_K\|}{\sigma_{\min}(\Sigma_K)} - \eta \|\Sigma_{K'}\| \|R\| \right) \operatorname{tr}(\nabla_K J(K)^\top \nabla_K J(K)) \\ &\geq 4\eta \frac{\sigma_{\min}(R)\sigma_{\min}^2(DD^\top)}{\|\Sigma_{K^*}\|} \left(1 - \frac{\|\Sigma_{K'} - \Sigma_K\|}{\sigma_{\min}(DD^\top)} - \eta \|\Sigma_{K'}\| \|R\| \right) (J(K) - J(K^*)).\end{aligned}$$

The condition on η ensures

$$\|K' - K\| \le \frac{\sigma_{\min}(Q)\sigma_{\min}(DD^\top)}{4J(K)\|B\|},$$

so by Lemma A.5,

$$\frac{\|\Sigma_{K'} - \Sigma_K\|}{\sigma_{\min}(DD^\top)} \le 4\eta \left(\frac{J(K)}{\sigma_{\min}(Q)\sigma_{\min}(DD^\top)}\right)^2 \|B\|\|\nabla_K J(K)\| \le \frac{1}{4},$$

with the assumed η. Then

$$\|\Sigma_{K'}\| \le \|\Sigma_K\| + \|\Sigma_{K'} - \Sigma_K\| \le \frac{J(K)}{\sigma_{\min}(Q)} + \frac{\sigma_{\min}(DD^\top)}{4} \le \frac{J(K)}{\sigma_{\min}(Q)} + \frac{\|\Sigma_{K'}\|}{4},$$

which implies $\|\Sigma_{K'}\| \le \frac{4J(K)}{3\sigma_{\min}(Q)}$. Hence,

$$1 - \frac{\|\Sigma_{K'} - \Sigma_K\|}{\sigma_{\min}(DD^\top)} - \eta\|\Sigma_{K'}\|\|R\| \ge 1 - \frac{1}{4} - \eta\frac{4J(K)\|R\|}{3\sigma_{\min}(Q)} \ge \frac{1}{4},$$

with the assumed η. Now we have

$$J(K) - J(K') \ge \eta\frac{\sigma_{\min}(R)\sigma_{\min}^2(DD^\top)}{\|\Sigma_{K^*}\|}(J(K) - J(K^*)),$$

which is equivalent to the desired conclusion. □

Theorem A.7. (Theorem 3). With an appropriate constant setting of the stepsize η in the form of

$$\eta = \text{poly}\left(\frac{\sigma_{\min}(Q)}{C(K_0)}, \sigma_{\min}(DD^\top), \frac{1}{\|B\|}, \frac{1}{\|R\|}\right),$$

and number of iterations

$$N \ge \frac{\|\Sigma_{K^*}\|}{\eta\sigma_{\min}^2(DD^\top)\sigma_{\min}(R)} \log\frac{J(K_0) - J(K^*)}{\varepsilon},$$

the iterates of gradient descent enjoys

$$J(K_N) - J(K^*) \le \varepsilon.$$

Proof. Iterating the gradient decent for N times, from Lemma A.6, we know

$$J(K_N) - J(K^*) \le \left(1 - \eta\frac{\sigma_{\min}(R)\sigma_{\min}^2(DD^\top)}{\|\Sigma_{K^*}\|}\right)^N (J(K_0) - J(K^*)).$$

Therefore, if N is chosen as the above, we can make the right hand side smaller than ε. □

B Proofs for Section 4

Proposition B.1. (Proposition 4). Assume $A - BK$ is stable. The optimal intercept b^K to minimize $J_2(K, b)$ for any given K is that

$$b^K = -(KQ^{-1}A^\top + R^{-1}B^\top)(AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}a \tag{B.1}$$

Furthermore, $J_2(K, b^K)$ takes the form of

$$J_2(K, b^K) = a^\top(AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}a \tag{B.2}$$

which is independent of K.

Proof. The problem of $\min_b J_2(K,b)$ is equivalent to the following constrained optimization

$$\begin{aligned} &\min \begin{pmatrix} \mu \\ b \end{pmatrix}^\top \begin{pmatrix} Q + K^\top RK & -K^\top R \\ -RK & R \end{pmatrix} \begin{pmatrix} \mu \\ b \end{pmatrix} \\ &\text{s.t. } (A - BK)\mu + (a + Bb) = 0 \end{aligned} \tag{B.3}$$

Using the Lagrangian multiplier method, we have

$$2M \begin{pmatrix} \mu \\ b \end{pmatrix} + N\lambda = 0 \,, \qquad N^\top \begin{pmatrix} \mu \\ b \end{pmatrix} + a = 0 \,,$$

where

$$M = \begin{pmatrix} Q + K^\top RK & -K^\top R \\ -RK & R \end{pmatrix}, \qquad N = \begin{pmatrix} (A - BK)^\top \\ B^\top \end{pmatrix}.$$

From the first equation we get $(\mu^\top, b^\top)^\top = -M^{-1}N\lambda/2$. Plugging this into the second equation, we derive $\lambda = -2(N^\top M^{-1}N)^{-1}a$. Therefore, the optimal (μ^K, b^K) is

$$\begin{pmatrix} \mu^K \\ b^K \end{pmatrix} = -M^{-1}N(N^\top M^{-1}N)^{-1}a \,.$$

And the optimal value of $J_2(K,b)$ is $J_2(K,b^K) = a^\top(N^\top M^{-1}N)^{-1}a$. By some simple calculation,

$$M^{-1} = \begin{pmatrix} Q^{-1} & -Q^{-1}K^\top \\ -KQ^{-1} & KQ^{-1}K^\top + R^{-1} \end{pmatrix},$$

and $N^\top M^{-1}N = AQ^{-1}A^\top + BR^{-1}B^\top$. Therefore, the final optimal

$$\begin{pmatrix} \mu^K \\ b^K \end{pmatrix} = -\begin{pmatrix} Q^{-1}A^\top \\ KQ^{-1}A^\top + R^{-1}B^\top \end{pmatrix} (AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}a \,.$$

We have assumed M and $N^\top M^{-1}N$ are non-singular above. We now rigorously show that they are indeed invertible. Specifically, if M is singular, $\exists x = (x_1^\top, x_2^\top)^\top \neq 0$ but $x^\top Mx = 0$. Since $Q \succ 0$, we have $x_1 = 0$. Since $R \succ 0$, we have $-Kx_1 + x_2 = 0$, thus $x_2 = 0$. Then we get a contradiction. If $N^\top M^{-1}N$ is singular, $\exists x \neq 0$, but $Nx = 0$, which leads to $(A - BK)x = 0$. Given that $A - BK$ is stable, this implies $x = 0$, again we get a contradiction. The proof is now complete.

□

Theorem B.2. (Theorem 5). With the stepsize η in the form of

$$\eta = \text{poly}\left(\frac{\sigma_{\min}(Q)}{C(K_0)}, \sigma_{\min}(DD^\top), \frac{1}{\|B\|}, \frac{1}{\|R\|} \right),$$

and number of iterations

$$N \geq \frac{\|\Sigma_{K^*}\|}{\eta\sigma_{\min}^2(DD^\top)\sigma_{\min}(R)} \log \frac{J_1(K_0) - J_1(K^*)}{\varepsilon},$$

the iterates of gradient descent enjoys $J_1(K_N) - J_1(K^*) \leq \varepsilon$. If we follow $b^K = -(KQ^{-1}A^\top + R^{-1}B^\top)(AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}a$, we have

$$J(K_N, b^{K_N}) - J(K^*, b^*) \leq \varepsilon.$$

Furthermore,

$$\|K_N - K^*\|_F \leq \sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon}, \quad \|b^{K_N} - b^*\|_2 \leq C_b(a)\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon}\,, \tag{B.4}$$

where $C_b(a) = \|Q^{-1}A^\top(AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}a\|_2$ is a constant depending on the intercept a.

Proof. We only need to show the bound for K_N and b^{K_N} in (B.4). From the proof of Lemma 2, we showed that for any K, K',

$$J_1(K) - J_1(K') = \mathrm{tr}[\Sigma_K[E_{K'}^\top (K-K') + (K-K')^\top E_{K'} + (K-K')^\top R(K-K')]] \,.$$

Choosing $K' = K^*$, since $E_{K^*} = 0$, we get

$$J_1(K) - J_1(K^*) = \mathrm{tr}[\Sigma_K (K-K^*)^\top R(K-K^*)] \geq \sigma_{\min}(R), \sigma_{\min}(DD^\top)\|K_N - K^*\|_F^2 \,.$$

Therefore, if (K_N, b^{K_N}) makes $J(K_N, b^{K_N}) - J(K^*, b^*) = J_1(K) - J_1(K^*) \leq \varepsilon$, we surely obtain $\|K_N - K^*\|_F^2 \leq \sigma_{\min}^{-1}(R)\sigma_{\min}^{-1}(DD^\top)\varepsilon$.

The bound for b^{K_N} is straightforward as

$$\begin{aligned}\|b^{K_N} - b^*\|_2 &\leq \|K_N - K^*\|_2 \|Q^{-1}A^\top (AQ^{-1}A^\top + BR^{-1}B^\top)^{-1} a\|_2 \\ &\leq C_b(a)\|K_N - K^*\|_F \leq C_b(a)\ \sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon}\,.\end{aligned}$$

□

C Proofs for Section 5

Proposition C.1. (Proposition 8). Under Assumption 7, the operator $\Lambda(\cdot) = \Lambda_2(\cdot, \Lambda_1(\cdot))$ is L_0-Lipschitz, where L_0 is given in Assumption 7. Moreover, there exists a unique Nash equilibrium pair (μ^*, π^*) of the MFG.

Proof. Consider the linear policies $\pi_{K,b}(x) = -Kx + b$. Define the distance metric of the linear policy as follows

$$d(\pi_{K_1,b_1}, \pi_{K_2,b_2}) = \|K_1 - K_2\|_2 + \|b_1 - b_2\|_2 \,. \tag{C.1}$$

Then for the mapping $\Lambda_1(\mu)$, as the optimal K^* does not depend on μ, we have for any $\mu_1, \mu_2 \in \mathbb{R}^{d+k}$,

$$\begin{aligned} d(\Lambda_1(\mu_1), \Lambda_2(\mu_2)) &= \|b_{1,\mu}^* - b_{2,\mu}^*\|_2 \\ &\leq \|K^*Q^{-1}A^\top + R^{-1}B^\top\|_2 \Big(\Big\|(AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}\bar{A}\Big\|_2 \|\mu_{1,x} - \mu_{2,x}\|_2 \\ &\quad + \Big\|(AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}\bar{B}\Big\|_2 \|\mu_{1,u} - \mu_{2,u}\|_2 \Big) \\ &\leq L_1(\|\mu_{1,x} - \mu_{2,x}\|_2 + \|\mu_{1,u} - \mu_{2,u}\|_2) = L_1\|\mu_1 - \mu_2\|_2 \,. \end{aligned} \tag{C.2}$$

For the mapping $\Lambda_2(\mu, \pi)$, with the same optimal policy $\pi \in \Pi$ under some $\mu \in \mathbb{R}^{d+k}$, for any $\mu_1, \mu_2 \in \mathbb{R}^{d+k}$, it holds that

$$\begin{aligned} \|\Lambda_2(\mu_1, \pi) - \Lambda_2(\mu_2, \pi)\|_2 &= \|\mu_{\mathrm{new},x}(\mu_1) - \mu_{\mathrm{new},x}(\mu_2)\|_2 + \|\mu_{\mathrm{new},u}(\mu_1) - \mu_{\mathrm{new},u}(\mu_2)\|_2 \\ &\leq \|(A - BK^*)^{-1}\bar{A}\|_2 \|\mu_{1,x} - \mu_{2,x}\|_2 \\ &\quad + \|(A - BK^*)^{-1}\bar{B}\|_2 \|\mu_{1,u} - \mu_{2,u}\|_2 \\ &\quad + \|K^*(A - BK^*)^{-1}\bar{A}\|_2 \|\mu_{1,x} - \mu_{2,x}\|_2 \\ &\quad + \|K^*(A - BK^*)^{-1}\bar{B}\|_2 \|\mu_{1,u} - \mu_{2,u}\|_2 \\ &\leq L_2(\|\mu_{1,x} - \mu_{2,x}\|_2 + \|\mu_{1,u} - \mu_{2,u}\|_2) = L_2\|\mu_1 - \mu_2\|_2 \,. \end{aligned} \tag{C.3}$$

With the same mean-field variable μ, since any two optimal policies π_1 and π_2 share the same K^*, we also have the following bound

$$\|\Lambda_2(\mu,\pi_1)-\Lambda_2(\mu,\pi_2)\|_2 \le \Big(\|(A-BK^*)^{-1}B\|_2+\|I+K^*(A-BK^*)^{-1}B\|_2\Big)\|b_{\pi_1}-b_{\pi_2}\|_2 = L_3\|b_{\pi_1}-b_{\pi_2}\|_2\,. \tag{C.4}$$

Therefore, combining (C.2). (C.3), (C.4), we obtain for any $\mu_1,\mu_2\in\mathbb{R}^{d+k}$,

$$\begin{aligned}\|\Lambda(\mu_1)-\Lambda(\mu_2)\|_2 &= \|\Lambda_2(\mu_1,\Lambda_1(\mu_1))-\Lambda_2(\mu_2,\Lambda_1(\mu_2))\|_2\\ &\le \|\Lambda_2(\mu_1,\Lambda_1(\mu_1))-\Lambda_2(\mu_1,\Lambda_1(\mu_2))\|_2+\|\Lambda_2(\mu_1,\Lambda_1(\mu_2))-\Lambda_2(\mu_2,\Lambda_1(\mu_2))\|_2\\ &\le L_3\ d(\Lambda_1(\mu_1),\Lambda_1(\mu_2))+L_2\|\mu_1-\mu_2\|_2\\ &\le (L_1L_3+L_2)\ \|\mu_1-\mu_2\|_2 = L_0\,\|\mu_1-\mu_2\|_2\,.\end{aligned} \tag{C.5}$$

So given the assumption that $L_0<1$, the operator $\Lambda(\cdot)$ is a contraction. By Banach fixed-point theorem, we conclude that $\Lambda(\cdot)$ has a unique fixed point, which gives the unique Nash equilibrium pair. This completes the proof of the proposition.

□

Theorem C.2. (Theorem 9). For a sufficiently small tolerance $0<\varepsilon<1$, we choose the number of iterations S in Algorithm 1 such that

$$S \ge \frac{\log(2\|\mu_0-\mu^*\|_2\cdot\varepsilon^{-1})}{\log(1/L_0)}\,. \tag{C.6}$$

For any $s=0,1,\ldots,S-1$, define

$$\begin{aligned}\varepsilon_s = \min\Big\{&2^{-2}\|B\|_2^{-2}\|(A-BK^*)^{-1}\|_2^{-2}, C_b(\mu_s)^{-2}\varepsilon^2, &\text{(C.7)}\\ &2^{-2s-4}(L_3C_b(\mu_s)+2C_K(\mu_2))^{-2}\varepsilon^2,\varepsilon^2\Big\}\cdot\sigma_{\min}(R)\sigma_{\min}(DD^\top)\,, &\text{(C.8)}\end{aligned}$$

where

$$\begin{aligned}C_b(\mu_s) &= \|Q^{-1}A^\top(AQ^{-1}A^\top+BR^{-1}B^\top)^{-1}\widetilde{a}_{\mu_s}\|_2\,, &\text{(C.9)}\\ C_K(\mu_s) &= \Big(\|\widetilde{\alpha}_{\mu_s}\|_2+(1+L_1\|\mu_s\|_2)\|B\|_2\Big)\\ &\quad\cdot\Big(\|(A-BK^*)^{-1}\|_2+(1+\|K^*\|_2)\|(A-BK^*)^{-1}\|_2^2\|B\|_2\Big)\,. &\text{(C.10)}\end{aligned}$$

In the s-th policy update, we choose the stepsize η as in Theorem 5 and number of iterations

$$N_s \ge \frac{\|\Sigma_{K^*}\|}{\eta\sigma_{\min}^2(DD^\top)\sigma_{\min}(R)}\log\frac{J_{\mu_s,1}(K_{\pi_s})-J_{\mu_s,1}(K^*)}{\varepsilon_s},$$

such that $J_{\mu_s}(K_{\pi_{s+1}},b_{\pi_{s+1}})-J_{\mu_s}(K^*,b^*_{\mu_s})\le\varepsilon_s$ where $K^*,b^*_{\mu_s}$ are parameters of the optimal policy $\pi^*_{\mu_s}=\Lambda_1(\mu_s)$ generated from mean-field state/action μ_s, $J_{\mu_s}(K_\pi,b_\pi)=J_{\mu_s}(\pi)$ is defined in the drifted MFG problem (17), and $J_{\mu_s,1}(K_\pi)$ is defined in (14) corresponding to $J_{\mu_s}(K_\pi,b_\pi)$. Then it holds that

$$\|\mu_S-\mu^*\|_2\le\varepsilon,\qquad \|K_{\pi_S}-K^*\|_F\le\varepsilon,\qquad \|b_{\pi_S}-b^*\|_2\le(1+L_1)\varepsilon. \tag{C.11}$$

Here μ^* is the Hash mean-field state/action, K_{π_S},b_{π_S} are parameters of the final output policy π_S, and K^*,b^* are the parameteris of the Nash policy $\pi^*=\Lambda_1(\mu^*)$.

Proof. Define $\mu_{s+1}^* = \Lambda(\mu_s)$ as the mean-field state/action generated by the optimal policy $\pi_{\mu_s}^* = \Lambda_1(\mu_s)$. Then by (19) and (20), we know that $\mu_{s+1}^* = ({\mu_{s+1,x}^*}^\top, {\mu_{s+1,u}^*}^\top)^\top$, and

$$\mu_{s+1,x}^* = -(A - BK^*)^{-1}(Bb_{\mu_s}^* + \widetilde{\alpha}_{\mu_s}),$$
$$\mu_{s+1,u}^* = b_{\mu_s}^* + K^*(A - BK^*)^{-1}(Bb_{\mu_s}^* + \widetilde{\alpha}_{\mu_s}).$$

Therefore, by triangle inequality,

$$\|\mu_{s+1} - \mu^*\|_2 \le \|\mu_{s+1} - \mu_{s+1}^*\|_2 + \|\mu_{s+1}^* - \mu^*\|_2 = E_1 + E_2. \tag{C.12}$$

Next we bound E_1 and E_2 separately.

The bound for E_2 is straighforward. From Proposition 8, we have

$$E_2 = \|\mu_{s+1}^* - \mu^*\|_2 = \|\Lambda(\mu_s) - \Lambda(\mu^*)\|_2 \le L_0\|\mu_s^* - \mu^*\|_2,$$

where $L_0 = L_1L_3 + L_2$ is defined in Assumption 7.

The bound for E_1 is more involved.

$$\begin{aligned} E_1 &= \|\mu_{s+1} - \mu_{s+1}^*\|_2 = \|\mu_{s+1,x} - \mu_{s+1,x}^*\|_2 + \|\mu_{s+1,u} - \mu_{s+1,u}^*\|_2 \\ &\le \Big(\|(A - BK^*)^{-1}B\|_2 + \|I + K^*(A - BK^*)^{-1}B\|_2\Big)\|b_{\pi_{s+1}} - b_{\mu_s}^*\|_2 \\ &\quad + \|Bb_{\pi_{s+1}} + \widetilde{\alpha}_{\mu_s}\|_2\Big(\|(A - BK_{\pi_{s+1}})^{-1} - (A - BK^*)^{-1}\|_2 \\ &\quad + \|K_{\pi_{s+1}}(A - BK_{\pi_{s+1}})^{-1} - K^*(A - BK^*)^{-1}\|_2\Big) = F_1 + F_2. \end{aligned}$$

From Theorem 5, we have $\|b_{\pi_{s+1}} - b_{\mu_s}^*\|_2 \le C_b(\mu_s)\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s}$, where $C_b(\mu_s) = \|Q^{-1}A^\top(AQ^{-1}A^\top + BR^{-1}B^\top)^{-1}\widetilde{a}_{\mu_s}\|_2$. So

$$F_1 \le L_3C_b(\mu_s)\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s}. \tag{C.13}$$

Recall that $L_3 = \|(A - BK^*)^{-1}B\|_2 + \|I + K^*(A - BK^*)^{-1}B\|_2$ is defined in Assumption 7. Now let us bound F_2.

Firstly,

$$\begin{aligned} \|Bb_{\pi_{s+1}} + \widetilde{\alpha}_{\mu_s}\|_2 &\le \|Bb_{\mu_s}^* + \widetilde{\alpha}_{\mu_s}\|_2 + \|B\|_2\|b_{\pi_{s+1}} - b_{\mu_s}^*\|_2 \\ &\le (\|\widetilde{\alpha}_{\mu_s}\|_2 + L_1\|B\|_2\|\mu_s\|_2) + \|B\|_2C_b(\mu_s)\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s} \\ &\le \|\widetilde{\alpha}_{\mu_s}\|_2 + (L_1\|\mu_s\|_2 + 1)\|B\|_2, \end{aligned}$$

if we choose ε_s such that $C_b(\mu_s)\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s} \le 1$. The second inequality is due to L_1-Lipschitz of $\Lambda_1(\cdot)$. Secondly,

$$\|(A - BK_{\pi_{s+1}})^{-1} - (A - BK^*)^{-1}\|_2 \le \|(A - BK_{\pi_{s+1}})^{-1}\|_2\|(A - BK^*)^{-1}\|_2\|B(K_{\pi_{s+1}} - K^*)\|_2.$$

Therefore,

$$\begin{aligned} \|(A - BK_{\pi_{s+1}})^{-1} - (A - BK^*)^{-1}\|_2 &\le \frac{\|(A - BK^*)^{-1}\|_2^2\|B\|_2\|K_{\pi_{s+1}} - K^*\|_2}{1 - \|(A - BK^*)^{-1}\|_2\|B\|_2\|K_{\pi_{s+1}} - K^*\|_2} \\ &\le 2\|(A - BK^*)^{-1}\|_2^2\|B\|_2\|K_{\pi_{s+1}} - K^*\|_2, \end{aligned}$$

if we choose ε_s such that $\|(A-BK^*)^{-1}\|_2\|B\|_2\|K_{\pi_{s+1}}-K^*\|_2 \le \|(A-BK^*)^{-1}\|_2\|B\|_2\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s} \le 1/2$ where we use the bound $\|K_{\pi_{s+1}} - K^*\|_2 \le \sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s}$ from Theorem 5. Lastly,

$$
\begin{aligned}
&\|K_{\pi_{s+1}}(A - BK_{\pi_{s+1}})^{-1} - K^*(A - BK^*)^{-1}\|_2 \\
&\quad \le \|K_{\pi_{s+1}} - K^*\|_2\|(A - BK_{\pi_{s+1}})^{-1}\|_2 + \|K^*\|_2\|(A - BK_{\pi_{s+1}})^{-1} - (A - BK^*)^{-1}\|_2 \\
&\quad \le \|K_{\pi_{s+1}} - K^*\|_2\|(A - BK_{\pi_{s+1}})^{-1}\|_2 + 2\|K^*\|_2\|(A - BK^*)^{-1}\|_2^2\|B\|_2\|K_{\pi_{s+1}} - K^*\|_2 \\
&\quad \le 2\|K_{\pi_{s+1}} - K^*\|_2\|(A - BK^*)^{-1}\|_2 + 2\|K^*\|_2\|(A - BK^*)^{-1}\|_2^2\|B\|_2\|K_{\pi_{s+1}} - K^*\|_2 \,,
\end{aligned}
$$

where the last inequality assumes $\|(A - BK^*)^{-1}\|_2\|B\|_2\|K_{\pi_{s+1}} - K^*\|_2 \le 1/2$ again. Combing the above derivations, we reach the following bound for F_2

$$
F_2 \le 2C_K(\mu_s)\|K_{\pi_{s+1}} - K^*\|_2 \le 2C_K(\mu_s)\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s}\,, \tag{C.14}
$$

where

$$
C_K(\mu_s) = \Big(\|\widetilde{\alpha}_{\mu_s}\|_2 + (1 + L_1\|\mu_s\|_2)\|B\|_2\Big)\Big(\|(A - BK^*)^{-1}\|_2 + (1 + \|K^*\|_2)\|(A - BK^*)^{-1}\|_2^2\|B\|_2\Big)\,.
$$

Combining the bounds (C.13) and (C.14), we have

$$
E_1 \le (L_3C_b(\mu_s) + 2C_K(\mu_s))\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_s}\,.
$$

Finally, we hope to choose ε_s such that $E_1 \le \varepsilon \cdot 2^{-s-2}$, which will be sufficient to prove the theorem. Therefore, we just need to set ε_s as follows

$$
\begin{aligned}
\varepsilon_s = \min\Big\{&2^{-2}\|B\|_2^{-2}\|(A - BK^*)^{-1}\|_2^{-2}, C_b(\mu_s)^{-2}, \\
&2^{-2s-4}(L_3C_b(\mu_s) + 2C_K(\mu_2))^{-2}\varepsilon^2\Big\} \cdot \sigma_{\min}(R)\sigma_{\min}(DD^\top)\,.
\end{aligned}
$$

With the bounds of E_1 and E_2, we have shown from (C.12) that

$$
\|\mu_{s+1} - \mu^*\|_2 \le L_0\|\mu_s - \mu^*\|_2 + \varepsilon \cdot 2^{-s-2}\,. \tag{C.15}
$$

Iterating over s and noting that $L_0 < 1$, we have

$$
\|\mu_S - \mu^*\|_2 \le L_0^S\|\mu_0 - \mu^*\|_2 + \varepsilon/2\,.
$$

Therefore, if we choose $S > \log(2\|\mu_0 - \mu^*\|_2 \cdot \varepsilon^{-1})/\log(1/L_0)$, we have $\|\mu_S - \mu^*\|_2 < \varepsilon$.

Finally we show the bounds for K_{π_S} and b_{π_S}. Since K^* does not depend on μ_s, for any iteration s including the last iteration S, we directly get

$$
\|K_{\pi_S} - K^*\|_F \le \sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_S} \le \varepsilon\,, \tag{C.16}
$$

from Theorem 5. By the triangle inequality,

$$
\begin{aligned}
\|b_{\pi_S} - b^*\|_2 &\le \|b_{\pi_S} - b^*_{\mu_S}\|_2 + \|b^*_{\mu_S} - b^*\|_2 \\
&\le C_b(\mu_S)\sigma_{\min}^{-1/2}(R)\sigma_{\min}^{-1/2}(DD^\top)\sqrt{\varepsilon_S} + L_1\|\mu_S - \mu^*\|_2 \\
&\le (1 + L_1)\varepsilon\,,
\end{aligned} \tag{C.17}
$$

where the second inequality comes from Theorem 5 and the last inequality comes from the choice of ε_S. Thus we now complete the proof of the theorem.

□

References

[1] Panos J Antsaklis and Anthony N Michel. *Linear systems*. Springer Science & Business Media, 2006.

[2] Maryam Fazel, Rong Ge, Sham M Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator. *arXiv preprint arXiv:1801.05039*, 2018.

[3] Hamed Karimi, Julie Nutini, and Mark Schmidt. Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 795–811. Springer, 2016.

[4] Hesameddin Mohammadi, Armin Zare, Mahdi Soltanolkotabi, and Mihailo R Jovanović. Convergence and sample complexity of gradient methods for the model-free linear quadratic regulator problem. *arXiv preprint arXiv:1912.11899*, 2019.