SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Data-Driven Chebyshev Iteration for Linear Quadratic Gaussian Games

Yiheng Liu[1], Weinan Gao[1*] & Yongliang Yang[2]

[1]*State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China*
[2]*School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China*

Throughout this paper, $\mathbb{R}$ denotes the set of real numbers. The Kronecker product is represented by $\otimes$. For a symmetric matrix $P \in \mathbb{R}^{n\times n}$, vec$(P)$: the n^2-dimensional vector formed by stacking the columns of P on top of one another, $\beta(P) = \text{vec}(P) = [p_{1,1}, p_{1,2}, \ldots, p_{1,n}, \ldots, p_{n,1}, p_{n,2}, \ldots, p_{n,n}]^\top \in \mathbb{R}^{n^2}$. $P > (\geqslant) 0$ and $P < (\leqslant) 0$ indicates that the symmetric matrix P is positive definite (semidefinite) and negative definite (semidefinite), respectively. For $A \in \mathbb{R}^{n\times n}$, all the eigenvalues of A are denoted by $\lambda_i(A)$, $i = 1, 2, \ldots, n$, and the real part of each eigenvalue is denoted by $\text{Re}(\lambda_i(A))$, $i = 1, 2, \ldots, n$.

Appendix A Proof of Theorem 1

Assumption 1. The pair (A, B) is controllable and $Re(\lambda_i(A)) \neq \rho, i = 1, \ldots, n$.

Lemma A1. [1] Regarding the operators $\mathcal{L}$ and $\mathcal{H}$ defined in Definition 1, the following facts hold.

1. The matrices P^* and K^* satisfy $\mathcal{H}(K^*, P^*) = 0$, where $P^* = P^{*\top} > 0$ can be obtained uniquely by solving the ARE $0 = \mathcal{R}(P)$ and K^* can be determined as $K^* = \mathcal{K}(P^*)$ accordingly.

2. The matrices P_s^* and K_s^* satisfy $\mathcal{H}_f(K_s^*, P_s^*) = 0$, where $P_s^* = P_s{}^{*\top} > 0$ can be obtained uniquely by solving the ARE $0 = \mathcal{R}_f(P)$ and K_s^* can be determined as $K_s^* = \mathcal{K}(P_s^*)$ accordingly.

3. $\mathcal{L}(A_\rho, P) \leqslant (<) 0$ implies $P \geqslant (>) 0$, if A_ρ is Hurwitz.

Lemma A2. [2] $0 \geqslant \mathcal{R}(P_1) \geqslant \mathcal{R}(P_2)$ implies $P^* \leqslant P_1 \leqslant P_2$ for $P_i = P_i^T$ with $i = 1, 2$.

Proof.

1) We prove by mathematical induction.

i) We have $\mathcal{R}(\Phi_{0,0}) \leqslant 0$, $\mathcal{R}(\Psi_{0,0}) \leqslant 0$, $i = 0, s = 0$. First, the recursion (5) is identical to the Newton iteration, which can be rewritten as $\Phi_{0,1}(A_\rho + B\mathcal{K}(\Phi_{0,0})) + (A_\rho + B\mathcal{K}(\Phi_{0,0}))^\top \Phi_{0,1} + \mathcal{K}(\Phi_{0,0})^\top R\mathcal{K}(\Phi_{0,0}) + Q = 0$. Therefore, we can get $\Phi_{0,0} \geqslant \Phi_{0,1} \geqslant P_*$ based on [3]. The recursion (7) is also identical to the Newton iteration with the state weight matrix zero, which can be rewritten as $\Psi_{0,1}(A_\rho + \mathcal{K}(\Psi_{0,0})) + (A_\rho + \mathcal{K}(\Psi_{0,0}))^\top \Psi_{0,1} + \mathcal{K}(\Psi_{0,0})^\top R\mathcal{K}(\Psi_{0,0}) = 0$. Therefore, we can get $\Psi_{0,0} \geqslant \Psi_{0,1} \geqslant P_s^*$ based on [4]. Next, the recursion (6) is essentially a Chord iteration. It ensures $\mathcal{L}(A_\rho + B\mathcal{K}(\Phi_{0,0}), \Phi_{0,1} - \Phi_{0,2}) = \mathcal{R}(\Phi_{0,1}) \leqslant 0$. Considering the Hurwitz matrix $A_\rho + B\mathcal{K}(\Phi_{0,0})$ and Lemma A1, the above fact implies $\Phi_{0,0} \geqslant \Phi_{0,1} \geqslant \Phi_{0,2}$. Therefore, $\Phi_{0,0} - \Phi_{0,2} \geqslant \Phi_{0,0} - \Phi_{0,1} \geqslant 0$ and $(\Phi_{0,0} - \Phi_{0,1})BR^{-1}B^\top(\Phi_{0,0} - \Phi_{0,1}) \leqslant (\Phi_{0,0} - \Phi_{0,2})BR^{-1}B^\top(\Phi_{0,0} - \Phi_{0,2})$ which implies that

$$\Phi_{0,0}BR^{-1}B^\top(\Phi_{0,1} - \Phi_{0,2}) + (\Phi_{0,1} - \Phi_{0,2})BR^{-1}B^\top\Phi_{0,0} \geqslant \Phi_{0,0}BR^{-1}B^\top\Phi_{0,0} - \Phi_{0,2}BR^{-1}B^\top\Phi_{0,2}.$$

By adding $A_\rho^\top(\Phi_{0,2} - \Phi_{0,1})$ and $(\Phi_{0,2} - \Phi_{0,1})A_\rho$ to both sides, one has

$$\begin{aligned}\mathcal{R}'_{\Phi_{0,0}}(\Phi_{0,2} - \Phi_{0,1}) &= [A_\rho + BK(\Phi_{0,0})]^\top(\Phi_{0,2} - \Phi_{0,1}) + (\Phi_{0,2} - \Phi_{0,1})[A_\rho + BK(\Phi_{0,0})] \\ &\geqslant A_\rho^\top\Phi_{0,2} + \Phi_{0,2}A_\rho - \Phi_{0,2}BR^{-1}B^\top\Phi_{0,2} + Q - A_\rho^\top\Phi_{0,1} - \Phi_{0,1}A_\rho + \Phi_{0,1}BR^{-1}B^\top\Phi_{0,1} - Q \\ &= \mathcal{R}(\Phi_{0,2}) - \mathcal{R}(\Phi_{0,1}).\end{aligned} \tag{A1}$$

From (6) and (A1), one has $-\mathcal{R}(\Phi_{0,1}) \geqslant \mathcal{R}(\Phi_{0,2}) - \mathcal{R}(\Phi_{0,1})$, i.e., $\mathcal{R}(\Phi_{0,2}) \leqslant 0$. Therefore, we observe $\Phi_{0,2} \geqslant P^*$ according to Lemma A2. One can conclude that $\Phi_{0,0} \geqslant \Phi_{0,1} \geqslant \Phi_{0,2} = \Phi_{1,0} \geqslant P^*$. Apparently, $\Psi_{0,0} \geqslant \Psi_{0,1} \geqslant \Psi_{0,2} = \Psi_{1,0} \geqslant P_s^*$.

ii) Suppose $\mathcal{R}(\Phi_i) \leqslant 0$, $\mathcal{R}(\Psi_i) \leqslant 0$, and $i > 0, s > 0$. Let us show that $\mathcal{R}(\Phi_{i+1}) \leqslant 0$, $\mathcal{R}(\Psi_{i+1}) \leqslant 0$, $\Phi_{i,0} \geqslant \Phi_{i,1} \geqslant \Phi_{i,2} = \Phi_{i+1,0} \geqslant P^*$ and $\Psi_{s,0} \geqslant \Psi_{s,1} \geqslant \Psi_{s,2} = \Psi_{s+1,0} \geqslant P_s^*$. First, the recursion (5) is identical to the Newton iteration, which can be rewritten as $\Phi_{i,1}(A_\rho + B\mathcal{K}(\Phi_{i,0})) + (A_\rho + B\mathcal{K}(\Phi_{i,0}))^\top\Phi_{i,1} + \mathcal{K}(\Phi_{i,0})^\top R\mathcal{K}(\Phi_{i,0}) + Q = 0$. Therefore, we can get

* Corresponding author (email: gaown@mail.neu.edu.cn)

$\Phi_{i,0} \geqslant \Phi_{i,1} \geqslant P_*$ based on [3]. The recursion (7) is also identical to the Newton iteration with the state weight matrix as zero, which can be rewritten as $\Psi_{s,1}(A_\rho+\mathcal{K}(\Psi_{s,0}))+(A_\rho+\mathcal{K}(\Psi_{s,0}))^\top\Psi_{s,1}+\mathcal{K}(\Psi_{s,0})^\top R\mathcal{K}(\Psi_{s,0})=0$. Therefore, we can get $\Psi_{s,0} \geqslant \Psi_{s,1} \geqslant P_s^*$ based on [4]. Next, the recursion (6) is essentially a Chord iteration. It ensures $\mathcal{L}(A_\rho + B\mathcal{K}(\Phi_{i,0}), \Phi_{i,1} - \Phi_{i,2}) = \mathcal{R}(\Phi_{i,1}) \leqslant 0$. Considering the Hurwitz matrix $A_\rho + B\mathcal{K}(\Phi_{i,0})$ and Lemma A1, the above fact implies $\Phi_{i,0} \geqslant \Phi_{i,1} \geqslant \Phi_{i,2}$. Therefore, $\Phi_{i,0} - \Phi_{i,2} \geqslant \Phi_{i,0} - \Phi_{i,1} \geqslant 0$ and $(\Phi_{i,0} - \Phi_{i,1})BR^{-1}B^\top(\Phi_{i,0} - \Phi_{i,1}) \leqslant (\Phi_{i,0} - \Phi_{i,2})BR^{-1}B^\top(\Phi_{i,0} - \Phi_{i,2})$ which implies that $\Phi_{i,0}BR^{-1}B^\top(\Phi_{i,1} - \Phi_{i,2}) + (\Phi_{i,1} - \Phi_{i,2})BR^{-1}B^\top\Phi_{i,0} \geqslant \Phi_{i,0}BR^{-1}B^\top\Phi_{i,0} - \Phi_{i,2}BR^{-1}B^\top\Phi_{i,2}$. By adding $A_\rho^\top(\Phi_{i,2} - \Phi_{i,1})$ and $(\Phi_{i,2} - \Phi_{i,1})A_\rho$ to both sides, one has

$$\begin{aligned}\mathcal{R}'_{\Phi_{i,0}}(\Phi_{i,2} - \Phi_{i,1}) &= [A_\rho + BK(\Phi_{i,0})]^\top(\Phi_{i,2} - \Phi_{i,1}) + (\Phi_{i,2} - \Phi_{i,1})[A_\rho + BK(\Phi_{i,0})]\\ &\geqslant A_\rho^\top\Phi_{i,2} + \Phi_{i,2}A_\rho - \Phi_{i,2}BR^{-1}B^\top\Phi_{i,2} + Q - A_\rho^\top\Phi_{i,1} - \Phi_{i,1}A_\rho + \Phi_{i,1}BR^{-1}B^\top\Phi_{i,1} - Q\\ &= \mathcal{R}(\Phi_{i,2}) - \mathcal{R}(\Phi_{i,1}).\end{aligned} \tag{A2}$$

From (6) and (A2), one has $-\mathcal{R}(\Phi_{i,1}) \geqslant \mathcal{R}(\Phi_{i,2}) - \mathcal{R}(\Phi_{i,1})$, i.e., $\mathcal{R}(\Phi_{i,2}) \leqslant 0$. Therefore, we observe $\Phi_{i,2} \geqslant P^*$ according to Lemma A2. One can conclude that $\Phi_{i,0} \geqslant \Phi_{i,1} \geqslant \Phi_{i,2} = \Phi_{i+1,0} \geqslant P^*$. Apparently, $\Psi_{s,0} \geqslant \Psi_{s,1} \geqslant \Psi_{s,2} = \Psi_{s+1,0} \geqslant P_s^*$.

2) The fact $\Phi_i \geqslant \Phi_{i+1} \geqslant P^*$ implies $\lim_{i\to\infty}\Phi_i = \Phi_\infty$ and $\lim_{i\to\infty}\mathcal{K}(\Phi_i) = K_\infty$; See [3]. On the other hand, the recursion (5), (7) indicates that $\mathcal{R}(\Phi_\infty) = 0$. Recalling the fact that the ARE has a unique solution P^*, one can conclude that $\Phi_\infty = P^*$ and $\mathcal{K}(\Phi_\infty) = K^*$. similarly, $\lim_{s\to\infty}\Psi_s = \Psi_{s_\infty} = P_s^*$ and $\lim_{s\to\infty}\mathcal{K}(\Psi_s) = K_{s_\infty} = K_s^*$.

Appendix B The connection between Theorem 1 and the Chebyshev iteration

The Chebyshev Iteration algorithm has a periodically updated Fréchet derivative, which is composed of a Newton step and a Chord step. Recursion (5) and (7) are identical to the Newton iteration. To reduce computational burden, the Newton iteration for solving ARE does not update the Fréchet derivative after the first iteration, which is essentially the Chord iteration, as shown in recursion(6) and (8).

Appendix C The symbol definitions for (14) and (15)

$$\Theta_i(t,t+kT) = \begin{bmatrix}\bar\theta_i(t,t+T)\\ \vdots\\ \bar\theta_i(t+(k-1)T,t+kT)\end{bmatrix},\quad \Gamma_s(t,t+kT) = \begin{bmatrix}\bar\gamma_s(t,t+T)\\ \vdots\\ \bar\gamma_s(t+(k-1)T,t+kT)\end{bmatrix},$$

where $\bar\theta_i(t,t+T) = [\theta_x(t,t+T), \theta_i(t,t+T), \delta_t]$, $\bar\gamma_s(t,t+T) = [\theta_x(t,t+T), \gamma_s(t,t+T), \delta_t]$,

$\bar x_j = \left[([x_j]_1)^2, [x_j]_1[x_j]_2, \ldots, [x_j]_1[x_j]_n, [x_j]_2[x_j]_1, ([x_j]_2)^2, \ldots, [x_j]_2[x_j]_n, \ldots, [x_j]_n[x_j]_1, [x_j]_n[x_j]_2, \ldots, ([x_j]_n)^2\right]^\top$,

$\theta_x(t,t+T) = E\left[e^{-\rho(t+T)}\bar x_j(t+T) - e^{-\rho t}\bar x_j(t)\right]$, $\theta_i(t,t+T) = 2\theta_{xu}(t,t+T)\cdot(I_{n\times m}\otimes R) - 2\theta_{xx}(t,t+T)\cdot(I_{n\times n}\otimes\mathcal{K}(\Phi_i)^\top R)$,

$\gamma_s(t,t+T) = 2\theta_{xu}(t,t+T)\cdot(I_{n\times m}\otimes R) - 2\theta_{xx}(t,t+T)\cdot(I_{n\times n}\otimes\mathcal{K}(\Psi_s)^\top R)$, $\delta_t = e^{-\rho(t+T)} - e^{-\rho t}$,

$\theta_{xx}(t,t+T) = E\left[\int_t^{t+T} e^{-\rho\tau}x_j(\tau)\otimes x_j(\tau)d\tau\right]$, $\theta_{xu}(t,t+T) = E\left[\int_t^{t+T} e^{-\rho\tau}x_j(\tau)\otimes\alpha_j(\tau)d\tau\right]$.

$$\bar\varphi_{i,1}(t,t+kT) = \begin{bmatrix}\varphi_{i,1}(t,t+T)\\ \vdots\\ \varphi_{i,1}(t+(k-1)T,t+kT)\end{bmatrix},\quad \bar\psi_{s,1}(t,t+kT) = \begin{bmatrix}\psi_{s,1}(t,t+T)\\ \vdots\\ \psi_{s,1}(t+(k-1)T,t+kT)\end{bmatrix},$$

where $\varphi_{i,1}(t,t+T) = -\theta_{xx}(t,t+T)\cdot\mathrm{vec}\Big[Q+\mathcal{K}(\Phi_i)^\top R\mathcal{K}(\Phi_i)\Big]$, $\psi_{s,1}(t,t+T) = -\theta_{xx}(t,t+T)\cdot\mathrm{vec}\Big[\mathcal{K}(\Psi_s)^\top R\mathcal{K}(\Psi_s)\Big]$.

$$\xi_{i,1} = \begin{bmatrix}\beta(\Phi_{i,1})\\ \mathrm{vec}(\mathcal{K}(\Phi_{i,1}))\\ \theta_1\end{bmatrix}\in\mathbb{R}^r,\quad \zeta_{s,1} = \begin{bmatrix}\beta(\Psi_{s,1})\\ \mathrm{vec}(\mathcal{K}(\Psi_{s,1}))\\ \theta_2\end{bmatrix}\in\mathbb{R}^r,$$

where $r = n^2+mn+1$, $\theta_1 = \frac{1}{\rho}\mathrm{Tr}(DD^\top\Phi_{i,1})$, $\theta_2 = \frac{1}{\rho}\mathrm{Tr}(DD^\top\Psi_{s,1})$.

Appendix D The symbol definitions for (18) and (19)

$$\bar\varphi_{i,2}(t,t+kT) = \begin{bmatrix}\varphi_{i,2}(t,t+T)\\ \vdots\\ \varphi_{i,2}(t+(k-1)T,t+kT)\end{bmatrix},\quad \bar\psi_{s,2}(t,t+kT) = \begin{bmatrix}\psi_{s,2}(t,t+T)\\ \vdots\\ \psi_{s,2}(t+(k-1)T,t+kT)\end{bmatrix},$$

where $\varphi_{i,2}(t,t+T) = -\theta_{xx}(t,t+T)\cdot \mathrm{vec}\Big[Q+\mathcal{K}(\Phi_i)^\top R\mathcal{K}(\Phi_i)\Big] + \theta_{xx}(t,t+T)\cdot \mathrm{vec}\Big[D\big(\mathcal{K}(\Phi_i),\mathcal{K}(\Phi_{i,1})\big)\Big]$,
$\psi_{s,2}(t,t+T) = -\theta_{xx}(t,t+T)\cdot \mathrm{vec}\Big[\mathcal{K}(\Psi_s)^\top R\mathcal{K}(\Psi_s)\Big] + \theta_{xx}(t,t+T)\cdot \mathrm{vec}\Big[D\big(\mathcal{K}(\Psi_s),\mathcal{K}(\Psi_{s,1})\big)\Big]$.

$$\xi_{i,2} = \begin{bmatrix} \beta(\Phi_{i,2}) \\ \mathrm{vec}(\mathcal{K}(\Phi_{i,2})) \\ \theta_3 \end{bmatrix} \in \mathbb{R}^r, \quad \zeta_{s,2} = \begin{bmatrix} \beta(\Psi_{s,2}) \\ \mathrm{vec}(\mathcal{K}(\Psi_{s,2})) \\ \theta_4 \end{bmatrix} \in \mathbb{R}^r,$$

where $r = n^2 + mn + 1$, $\theta_3 = \frac{1}{\rho}\mathrm{Tr}(DD^\top \Phi_{i,2})$, $\theta_4 = \frac{1}{\rho}\mathrm{Tr}(DD^\top \Psi_{s,2})$.

Appendix E Proof of Theorem 2

Proof. In each iteration, starting from the stabilizing feedback gain K_i, the iterative gains $\mathcal{K}(\Phi_{i,1})$ and K_{i+1} can be uniquely determined, provided that the solutions $\Phi_{i,2}$, $\Psi_{s,2}$ to the Lyapunov equation (5)(6) exist. In addition, the matrices $\Phi_{i,1}$, $\Psi_{s,1}$ satisfy the LS equations (14)(15). The matrices $\Phi_{i,2}$, $\Psi_{s,2}$ satisfy the LS equations (18)(19), and Θ_i and Γ_s have the full column rank. By Theorem 1, the convergence of Algorithm 1 can be ensured.

Appendix F The construction purpose of Algorithm 1

Based on online data, effective cognition is carried out on mean-field models in dynamic uncertain environments. In scenarios where system matrices are unknown, intelligent online decisions are made relying on real-time data to solve AREs. Then, actions are formulated based on intelligent decisions, which are aligned with the so-called intelligent "cognition-decision-action" framework.

Appendix G Simulation example

In this section, we perform a numerical simulation for a large-population LQG games with 100 agents to validate the effectiveness of the proposed algorithm.

$$A = \begin{bmatrix} 5 & 3 \\ 10 & 12 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{bmatrix},$$

where $\lambda_1(A) = 2$ and $\lambda_2(A) = 15$, and there exists a matrix $K = -\begin{bmatrix} 35 & 25 \end{bmatrix}$ such that $A + BK$ is Hurwitz. w_j is a standard two-dimensional Brownian motion.

In this simulation, the parameters of the cost function (2) are given by $Q = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$, $R = 1$, and $\rho = 0.01$. Clearly, assumption 1 is verifiable. The following analytical solution is obtained by solving ARE:

$$\lim_{i\to\infty} \Phi_i = \begin{bmatrix} 232.2887 & 59.3007 \\ 59.3007 & 34.5712 \end{bmatrix}, \quad \lim_{i\to\infty} \mathcal{K}(\Phi_i) = \begin{bmatrix} 59.3007 & 34.5712 \end{bmatrix},$$
$$\lim_{s\to\infty} \Psi_s = \begin{bmatrix} 207.1523 & 56.5775 \\ 56.5775 & 33.9803 \end{bmatrix}, \quad \lim_{s\to\infty} \mathcal{K}(\Psi_s) = \begin{bmatrix} 56.5775 & 33.9803 \end{bmatrix}.$$

The simulation results for the PI model-free algorithm are sketched in Figs. G1-G2. As can be seen from Figs. G1-G2, the convergence is achieved at the fourth iteration. The simulation results for the Chebyshev iterative model-free algorithm are sketched in Figs. G3-G4 where one can observe that the convergence is achieved at the second iteration.

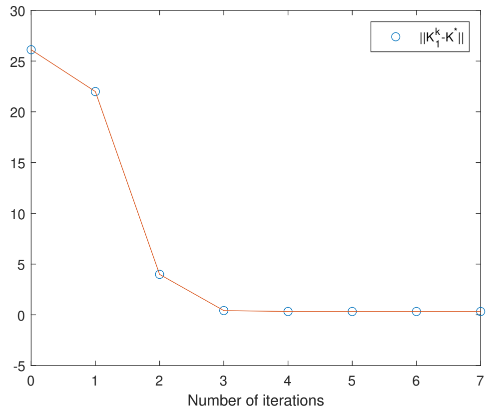

Figure G1 Evolution of Parameter of K^* in PI.

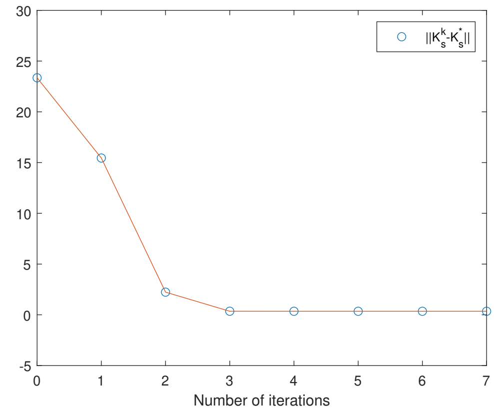

Figure G2 Evolution of Parameter of K_s^* in PI.

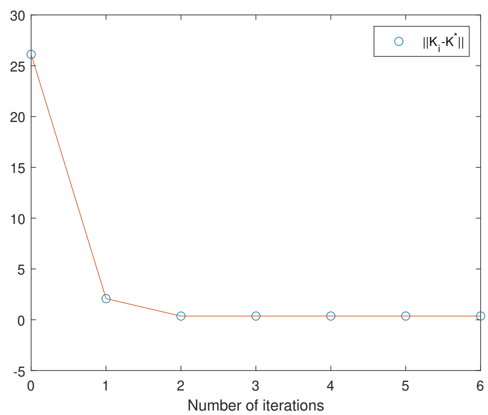

Figure G3 Evolution of Parameter of K^* in Chebyshev iteration.

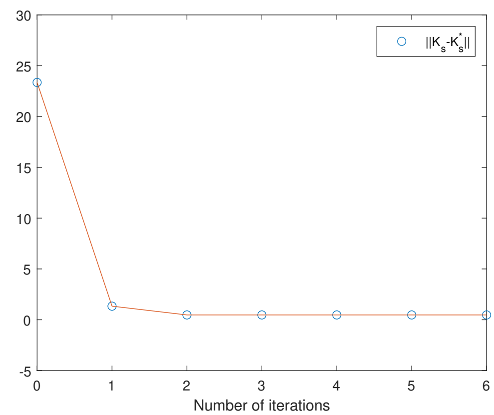

Figure G4 Evolution of Parameter of K_s^* in Chebyshev iteration.

References

1 Frank L. Lewis, Draguna L. Vrabie, Vassilis L. Syrmos. Optimal control. John Wiley & Sons, Inc, 2012.
2 Yang Y, Wunsch D, Yin Y. Hamiltonian-driven adaptive dynamic programming for continuous nonlinear dynamical systems. IEEE Trans Neural Netw Learn Syst, 2017, 28: 1929-1940
3 Kleinman D. On an iterative technique for Riccati equation computations. IEEE Trans Autom Control, 1968, 13: 114-115
4 Xu Z, Shen T, Huang M. Model-free policy iteration approach to NCE-based strategy design for linear quadratic Gaussian games. Automatica, 2023, 155: 111162