# Robust Linear Quadratic Reinforcement Learning

Ludvig Svedlund and Bengt Lennartson


## Extended Abstract

Knowledge about the dynamic behavior of a system is crucial in the design of feedback control systems. This knowledge is often based on system models, but also data from physical experiments. Such experiments can be step responses or time series analysis where the system is excited by some type of random input signals, often in combination with closed loop control. The controller is then adapted such that the behavior of the closed loop system is improved, naturally based on an online optimization procedure.

*Model-free reinforcement learning:* Reinforcement learning (RL) is a popular example of such an adaptive control strategy. Either a system model is then estimated, and based on this model a controller is designed. Alternatively, a controller is directly determined such that an estimated criterion is optimized. The first version is called model-based RL, while the second one is called model-free RL [1], [2], [3]. A special form of the model-free version is a state feedback controller that is optimized given a traditional linear quadratic criterion, among others formulated in [4], [5]. In this design strategy, the state feedback controller is determined based on experimental data, without knowing a state-space model for the system to be controlled. Including the correct state variables, the solution quickly converges to exactly the same solution as the one obtained by solving a Riccati equation based on a state-space model of the system [6].

*Low pass filtering when neglected dynamics is involved:* The focus of this paper is to evaluate what happens when this linear quadratic RL (LQRL) strategy is applied to a system where some dynamic behavior is neglected. Focusing on mechatronic systems, a second-order DC motor is investigated as the nominal model with an additional resonance or time constant as neglected dynamics. When the LQRL strategy only includes feedback from the two nominal states (motor angle and angular velocity), the learning procedure and closed loop stability are shown to be surprisingly sensitive to the unmodeled dynamics. Fortunately, low pass filtering of the control and nominal state signals is shown to significantly reduce the sensitivity to the unmodeled dynamics.

*Model-based reinforcement learning:* An interesting alternative is to evaluate the robustness to unmodeled dynamics in a corresponding model-based RL strategy. An ordinary least squares estimation of a second order state-space model is then performed based on the two nominal states and the control input signal. This strategy also works well without filtering, but the model-free LQRL version, with significant low pass filtering, sometimes obtains slightly better closed loop performance compared to the model-based strategy.

The model-based LQRL version also has some other benefits. The method is modularized in the sense that a system model is first estimated, which can be easily evaluated based on simple online experiments, such as step responses, but also based on physical knowledge of expected dynamic behavior. Model-free LQRL generates the final controller in one step, more as a black-box strategy.

*Output feedback reinforcement learning:* Often not all states are measured and then output feedback control is an important and common alternative control strategy. Model-free LQRL is then also shown to be a possible solution [7], but it has a built-in high gain strategy also explained in this paper. This high-gain behavior is less attractive especially when the controlled system has higher-order dynamics. Estimating an input-output model in model-based RL is on the other hand more simple than estimating a state-space model, and any control design method can then be used, including simple PID controllers and more advanced input-output control strategies.

*Nonlinear model-based reinforcement learning:* Also very importantly, a nonlinear model can be estimated, still very often based on a linear parameter regression model. This estimated nonlinear model can be locally linearized and a complete nonlinear control strategy is easily achieved. Furthermore, the convergence rate in model-based RL is less sensitive to disturbances [8], and the sample complexity is much more efficient in model-based compared to model-free RL [9].

*Summary:* To summarize the main contributions of this paper, it highlights the fact that traditional model-free LQRL is very sensitive to unmodeled dynamics, which fortunately can be greatly improved by low-pass filtering. The results are based on a fair evaluation method, based on both closed loop performance, control activity, and critical stability margins. Comparing the filtered model-free RL approach with model-based RL, the sensitivity to unmodeled dynamics is rather similar for the two methods. Since the model-based RL strategy, however, has other benefits such as being easier adapted to output feedback and nonlinear dynamics, as well as having much less sample complexity, it is strongly recommended as a generic and robust adaptive control strategy.


L. Svedlund and B. Lennartson are both with the Division of Systems and Control, Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden (e-mail: ludvige@chalmers.se).

*Future research:* For future research, it is recommended to examine the impact when more complex systems and unmodeled dynamics are introduced. Especially, the introduction of nonlinearities in the model-based approach will be compared with model-free deep RL.

## References


[1] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[2] D. Bertsekas, *Reinforcement learning and optimal control*. Athena Scientific, 2019, vol. 1.

[3] L. Busoniu, R. Babuska, B. D. Schutter, and D. Ernst, *Reinforcement learning and dynamic programming using function approximators*. CRC Press, 2010.

[4] S. J. Bradtke, B. E. Ydstie, and A. G. Barto, "Adaptive linear quadratic control using policy iteration," in *Proceedings of 1994 American Control Conference-ACC'94*, vol. 3. IEEE, 1994, pp. 3475–3479.

[5] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers," *IEEE Control Systems Magazine*, vol. 32, no. 6, pp. 76–105, 2012.

[6] B. D. O. Anderson and J. B. Moore, *Optimal Control: Linear Quadratic Methods*. Dover Publications, 2014.

[7] F. L. Lewis and K. G. Vamvoudakis, "Reinforcement learning for partially observable dynamic processes: Adaptive dynamic programming using measured output data," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 1, pp. 14–25, 2011.

[8] F. A. Yaghmaie, F. Gustafsson, and L. Ljung, "Linear quadratic control using model-free reinforcement learning," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 737–752, 2022.

[9] S. Tu and B. Recht, "The gap between model-based and model-free methods on the linear quadratic regulator: An asymptotic viewpoint," in *Conference on Learning Theory*. PMLR, 2019, pp. 3036–3083.