Trajectory optimization for swing-up of multi-link cart-pole

Siro Corsi
Underactuated Robotics final project
MIT 6.8210, Spring 2023

***Abstract*—The present study focuses on the analysis of a cartpole system consisting of multiple pendulums. To achieve swing-up motion, trajectory optimization and Finite Horizon Linear Quadratic Regulator (LQR) techniques are employed. Additionally, the unstable fixed point is effectively stabilized using LQR, and the local stability is certified through Sum of Squares verification. Moreover, a comprehensive analysis of the system's controllability is conducted by investigating the influence of pendulum lengths on the region of attraction.**

I. Introduction

The Cart-Pole System is widely recognized as a representative example of an underactuated robotic system. It comprises a cart that can move along a rail and is actuated by a force, with a pendulum attached on top, capable of swinging freely. Various control approaches have been developed and employed in control classes to address the swing-up and stabilization of this system [1]. An intriguing variation of the traditional cartpole system involves the incorporation of multiple pendulums. These additional pendulums can be arranged in a stacked configuration on top of each other or directly mounted on the cart itself. In this project, our analysis focuses on the scenario where the pendulums are all connected to the cart. Figure 1 provides a visual representation of a cartpole system with two pendulums. This particular configuration is of interest due to its inherent controllability, except when two pendulums possess identical lengths. Notably, the behavior of the determinant of the controllability matrix as a function of the pendulum lengths is illustrated in Figure 2. When two pendulums share the same length, the determinant becomes zero, resulting in an uncontrollable system. Consequently, this project aims to explore the impact of this phenomenon on the region of attraction associated with an LQR controller.

We commence our analysis by developing a controller specifically designed for the swing-up motion of a generalized cartpole system with multiple pendulums, each having distinct lengths. To accomplish this, trajectory optimization techniques are employed. Subsequently, we investigate the region of attraction associated with an LQR controller for the simplified cartpole system comprising a single pendulum. Additionally, we examine how the relative lengths of a system with two pendulums influence the region of attraction. To achieve this, we utilize sum of squares programs, employing a modified version of the cost-to-go function as a Lyapunov function. Furthermore, we employ sampling-based and simulation-based methods in our analysis. Our findings indicate that the cost-to-go function can be effectively employed to certify the stability of the nonlinear system. However, it is important to note that the region of attraction certified using this Lyapunov function is relatively small in comparison to the actual region.

For the implementation of our study, we rely on Drake [2] as a comprehensive toolbox.

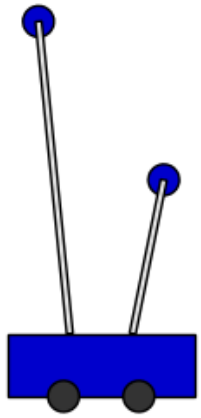

Fig. 1. An example of a cartpole with two pendulums.

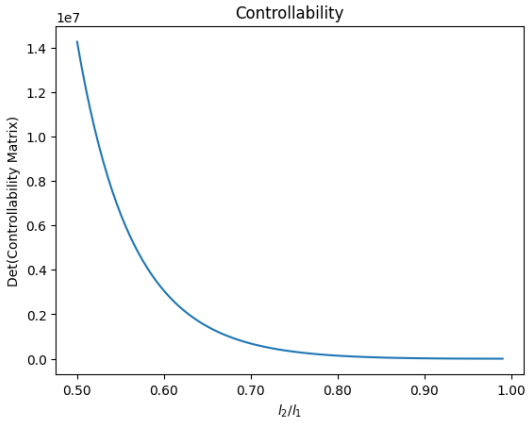

Fig. 2. The determinant of the controllability matrix goes to zero when $l_2 = l_1$.

II. Related work

The cartpole system has been extensively explored in the class textbook [1], primarily focusing on scenarios involving a single pendulum. The cartpole with two pendulums has been stabilyzed by a classical controller [4] (they call it dual-inverted-pendulum system).
In terms of swing-up techniques, energy-based methods have been applied to the cartpole system [6], [7]. Notably, Shen Shen successfully employed a sampling-based Sum of Squares verification method [5] to analyze an N-link pendulum mounted on a cart. This involved transforming the rational trigonometric equations of motion into polynomial form [3].

III. Method

A. Trajectory optimization

For the swing-up task, trajectory optimization, specifically direct collocation, is utilized. To initialize the trajectory, a linear interpolation is employed between the stable fixed point, where all pendulums are in a downward position, and the unstable fixed point, where the pendulums are in an upward position. The cost function employed in the optimization includes a quadratic term to penalize control efforts, and the final time to encourage faster swing-up. While input force limits are not considered, there are restrictions on the length of the rail. To stabilize the trajectory, a Finite Horizon Linear Quadratic Regulator is employed. Once all the pendulums are in the upright position, an Infinite Horizon LQR is applied for stabilization around the unstable fixed point.

B. Region of Attraction analysis

We conduct a comprehensive analysis of the region of attraction for the LQR controller around the unstable fixed point when all pendulums are in the upright position. Initially, we employ the cost-to-go function associated with the LQR controller as a Lyapunov function. Our objective is to identify a sublevel set where the time derivative is negative, thereby certifying the local stability of the closed-loop system. The sublevel set ensures that any initial condition within it will converge to the fixed point.

To determine the sublevel set, we employ Sum of Squares programs and sampling-based methods. To assess the efficacy of the cost-to-go function in certifying the region of attraction, we further validate the results through simulations.

1) Sum of Squares programs: The first Sum of Square program we use is the equality constrained formulation:

$$\begin{aligned} &\max_{\rho,\lambda(\mathbf{x})} \rho \\ &\text{subject to } \left(\mathbf{x}^T\mathbf{x}\right)^d (V(\mathbf{x})-\rho) + \lambda^T(\mathbf{x})\dot{V}(\mathbf{x}) \text{ is SOS,} \end{aligned} \tag{1}$$

This formulation necessitates that both V and $\dot{V}$ be polynomial, which is not the case for the cartpole due to its equations of motion being rational and involving trigonometric terms. However, we can introduce new variables, such as $s = \sin(\theta)$ and $c = \cos(\theta)$, to substitute the trigonometric terms, but with need to add the polynomial constraint $s^2+c^2-1=0$. To convert the equations from rational to polynomial form, the equations of motion can be kept in implicit form. Consequently, it is not possible to directly compute $\dot{V} = \frac{\partial V}{\partial x} f(x)$ (where $f(x)$ would be rational). Instead, the program should include the accelerations as indeterminates and incorporate the equations of motion in implicit form as constraints. This adjustment necessitates the use of additional multipliers to handle the equations of motion constraints. Additionally, modifications are required for the Lyapunov function and the force input. In the standard LQR formulation, both of these elements involve the angles θ, which cannot be included alongside the new variables s and c. However, since our focus is solely on local stability, we can substitute θ with s in the computation of both the cost-to-go and the control input. This substitution is acceptable since the linearization of θ and s around the fixed point yields the same result. This modified version of the cost-to-go function becomes polynomial, however it has the undesirable property of being 0 at the lower fixed point. To address this issue, we can add the term $1-c$ to the polynomial cost-to-go function. Additionally, careful consideration is required for the term $(\mathbf{x}^T\mathbf{x})$, as we cannot include θ in it. To handle this, we can use $\mathbf{x} = (x, s, c-1, \dot{x}, \dot{\theta})^T$, noting that at the fixed point, $c = 1$.

With these modifications, equation 1 can be utilized to find a value of ρ that certifies the stability of the closed-loop nonlinear system for every state $\mathbf{x}$ such that $V(\mathbf{x}) < \rho$.

Additionally, we utilize another Sum of Squares (SOS) formulation, as proposed by Shen Shen [5]. In this alternative formulation, instead of solving the SOS program for every state $\mathbf{x}$, it is claimed that it suffices to solve it for samples $\mathbf{x}_i$ where $\dot{V}(\mathbf{x}_i) = 0$. This formulation eliminates the need for multipliers and exhibits improved scalability compared to the equality constrained formulation.

2) Sampling based approaches: The introduction of the "Lyapunov analysis with convex optimization" chapter of the textbook [1] states "If you're imagining numerical algorithms to check the Lyapunov conditions for complicated Lyapunov functions and complicated dynamics, the first thought is probably that we can evaluate V and $\dot{V}$ at a large number of sample points and check whether V is positive and $\dot{V}$ is negative". We adopt a similar approach to estimate the largest sublevel set of the cost-to-go function, wherein we sample numerous points $\mathbf{x}_i$ and search for the minimum value of $V(\mathbf{x}_i)$ for which $\dot{V}(\mathbf{x}_i)$ is positive. This procedure provides an upper bound on the value of ρ.
To enhance the efficiency of sampling, we can further refine the process by sampling only $\mathbf{x}_i$ points where $\dot{V}(\mathbf{x}_i) = 0$. This approach minimizes the number of samples needed while still capturing the critical points of interest. Although the sampling process becomes more intricate, we can employ the same algorithm utilized for the sampling-based formulation of the SOS program. Remarkably, the required samples exhibit similar properties, enabling us to leverage the existing sampling algorithm for this purpose as well.

3) Simulation based approach: To assess the accuracy of each method in approximating the actual region of attraction, we conduct simulations using a significant number of initial conditions. We observe the system's behavior by monitoring whether the state converges towards the fixed point or diverges. This comparative analysis provides valuable insights into the effectiveness of each method in capturing the true region of attraction.

IV. Results and discussion

A. Swing up

Regarding the swing-up task, the results were promising. The direct collocation approach demonstrated efficiency and consistently produced satisfactory solutions across various configurations involving different numbers and lengths of pendulums. Notably, the project video [8] showcases specific examples highlighting the effectiveness of this method. However, it's worth mentioning that the number of collocation points needed to be adjusted for systems with varying numbers of pendulums, and the optimization time increased as more pendulums were added. While we successfully achieved swing-up control for up to 5 pendulums, we did not explore the limits beyond this threshold.

B. Region of attraction

We successfully obtained a solution for the equality constrained SOS program when dealing with the cartpole system with a single pendulum. However, when extending the system to multiple pendulums, the program's dimension became too large, and we were unable to find a viable solution. To address this issue, we turned to the sampling-based SOS formulation, which was expected to offer some improvements. Unfortunately, we encountered challenges in obtaining consistent and reliable results using this method. When using a small number of samples, the results appeared to be dependent on the specific set of samples chosen. In other words, different sample sets yielded different outcomes when solving the program. Increasing the number of samples did not alleviate the issue; instead, no solution could be found. We suspect that numerical difficulties were the primary obstacle. Despite our best efforts to carefully choose monomials and formulate the problem to guarantee the compatibility between the equality constrained program and the sampling-based SOS program, we encountered difficulties in finding a solution for the sampling-based program. Even though we were aware of the existence of a solution, we were unable to obtain it.

When applied to the single pendulum cartpole, both sampling-based approaches converged to the same value. As anticipated, the number of samples required for convergence was significantly lower when sampling points where $\dot{V}(\mathbf{x}_i) = 0$, as these samples were more informative. However, it is important to note that the sampling process itself was more computationally expensive. While we did not conduct an extensive evaluation comparing the convergence speed in terms of wall-clock time, it is worth mentioning that in our case, sampling points where $\dot{V}(\mathbf{x}_i) = 0$ resulted in faster convergence. The specific implementation and random factors can influence the relative speed, so a more comprehensive analysis is needed.

To visualize the actual region of attraction, we conducted simulations of the single pendulum cartpole system on a grid of initial conditions around the unstable fixed point, stabilized with LQR. Only the initial conditions of θ and $\dot{\theta}$ were modified, while x and $\dot{x}$ were set to 0. The results are presented in Figure 3. It is evident that the real region of attraction is significantly larger than what can be certified using the SOS approach, and even the sampling-based result is considerably smaller compared to the simulation-based result. This observation highlights that while it is possible to certify a limited region using a modified version of the cost-to-go, this Lyapunov function is conservative and underestimates the actual region of attraction.

In line with our introduction, where we highlighted the significance of the cartpole with multiple pendulums due to the change in controllability based on the pendulum lengths, we conducted simulations to examine how this variation impacts the region of attraction. Specifically, we explored the system's behavior across a range of initial conditions and pendulum lengths. Figure IV illustrates the relationship between the displacement of the initial angle for the longest and shortest pendulums respectively as the lengths of the two pendulums converge. Notably, we observe that the maximum stabilizable displacement decreases as the pendulum lengths approach each other. Furthermore, we observe a preference for displacements in the shorter pendulum rather than the longer one. This analysis provides valuable insights into the interplay between pendulum lengths and the resulting region of attraction for the system. It's important to note that all the visual representations provided, such as the two-dimensional plots, only capture a slice of the actual six-dimensional region of attraction. In these plots, we only consider the variations in a subset of the system's coordinates while keeping the remaining coordinates fixed at zero. In the project video [8], an animated version of the region of attraction is presented, providing a more dynamic visualization of its behavior. Figure IV showcases some frames from the animation. The video demonstrates how the region of attraction not only undergoes an expected shrinkage but also rotates. The rotation of the region is attributed to the influence of pendulum length disparities. When one pendulum is significantly shorter than the other, the LQR controller can effectively handle large displacements in the shorter link, resulting in a vertical orientation of the "ROA principal axis." Conversely, when the pendulums have the same length, a similar displacement is preferred, leading to the "ROA principal axis" being aligned at a 45-degree angle. In the video, it is possible to observe an unusual behavior in the region of attraction. However, this behavior is likely attributed to a discretization artifact rather than a fundamental characteristic of the system. Nevertheless, a comprehensive analysis has not been conducted to draw a definitive conclusion.

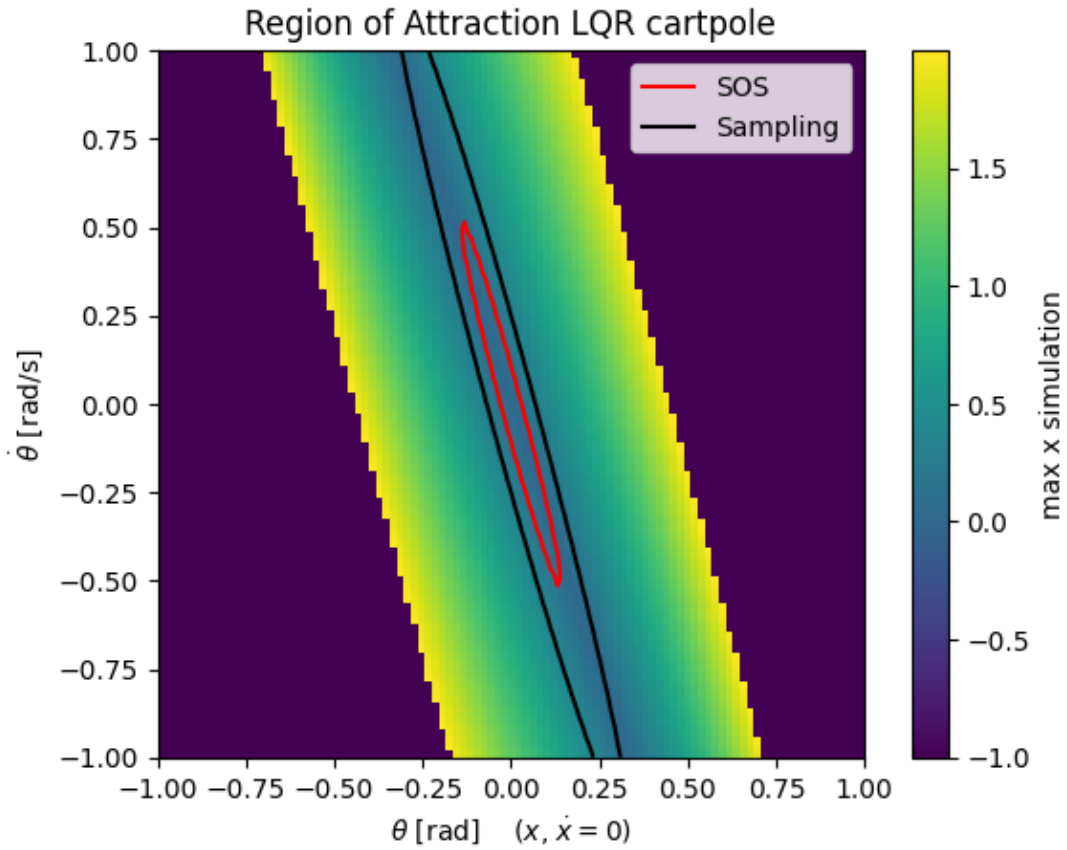

Fig. 3. Region of attraction for the single pendulum cartpole LQR controller. The color mapping represents the maximum displacement of the cart required to stabilize the pendulum. The red line depict the region certified by the SOS program, while the black line indicates the upper bound on ρ obtained through the sampling approach.

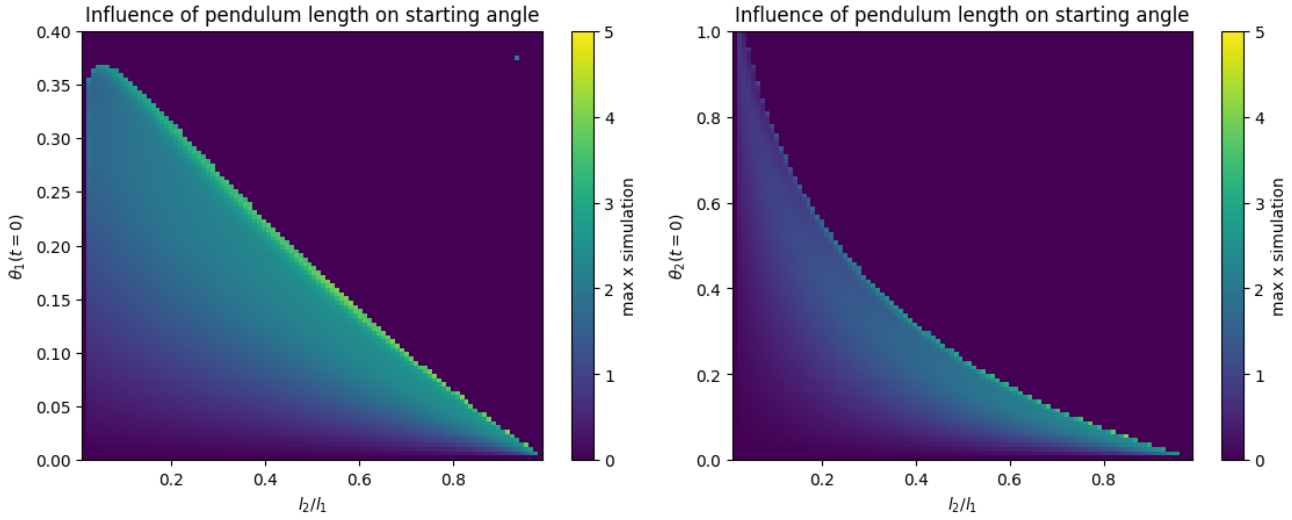

Fig. 4. The simulated region of attraction is depicted for two scenarios: a displacement in the longer pendulum (left) and a displacement in the shorter pendulum (right), varying with respect to the pendulum lengths. Please note the different scales used for each plot.

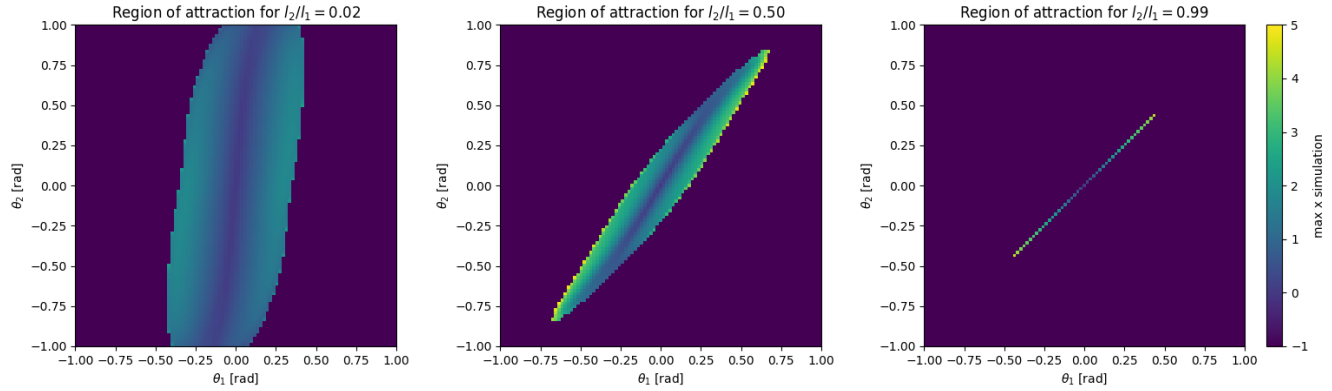

Fig. 5. Three frames from the animation showcased in the video [8] are presented. Each frame represents a slice of the region of attraction for a different value of l_2/l_1. The color mapping indicates the maximum distance the cart had to displace in order to stabilize the pendulums

V. Conclusion

In conclusion, this project provided an opportunity to apply various concepts covered in our course to a seemingly simple system. We successfully demonstrated the effectiveness of direct collocation in swinging up multiple pendulums on a cartpole. Additionally, we showcased how the LQR controller for the cartpole system can be certified using Sum of Squares programs with the cost-to-go serving as a Lyapunov function. Furthermore, we conducted an analysis of the region of attraction for the cartpole with two pendulums and observed its gradual shrinkage as the system approached the point of non-controllability.

References

[1] Russ Tedrake. Underactuated Robotics. Algorithms for Walking, Running, Swimming, Flying, and Manipulation. 2023. URL: http://underactuated.mit.edu.

[2] Russ Tedrake and the Drake Development Team. Drake: Model-based design and verification for robotics. 2019. URL: https://drake.mit.edu.

[3] A. Papachristodoulou and S. Prajna, “Analysis of Non-polynomial Systems Using the Sum of Squares Decomposition,” 2005.

[4] K. H. Lundberg and J. K. Roberge, ”Classical dual-inverted-pendulum control,” 42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475), Maui, HI, USA, 2003, pp. 4399-4404 Vol.5, doi: 10.1109/CDC.2003.1272198.

[5] S. Shen and R. Tedrake, ”Sampling Quotient-Ring Sum-of-Squares Programs for Scalable Verification of Nonlinear Systems,” 2020 59th IEEE Conference on Decision and Control (CDC), Jeju, Korea (South), 2020, pp. 2535-2542, doi: 10.1109/CDC42340.2020.9304028.

[6] : Xin, Xin & Kaneda, Masahiro. (2005). Analysis of the energy-based control for swing up two pendulums. Automatic Control, IEEE Transactions on. 50. 679 - 684. 10.1109/TAC.2005.846598.

[7] : Xin, Xin & Liu, Yannian. (2013). Control Design and Analysis for Underactuated Robotic Systems. 10.1007/978-1-4471-6251-3.

[8] : Project video: https://youtu.be/bhBG78Pbh6Y