

# Path Integral Stochastic Optimal Control for Dynamic System: Swing-up Task

Erivelton Gualter dos Santos<sup>†,\*</sup>, Hamed Kayello<sup>†</sup>

<sup>†</sup> Ph.D Candidate, Cleveland State University

<sup>‡</sup> Ph.D, Instructor, Cleveland State Univrsity

## 1. DYNAMIC REPRESENTATION OF STOCHASTIC SYSTEM

According to Morrison (2012), “*modeling is neither science nor mathematics; it is the craft that builds bridges between the two.*” It is an essential tool for mathematically describe a device’s behavior that can be computed and compared to the system’s measured information. In system dynamics and robotics, differential equations have significantly been used for this task and showed success in describing these behaviors (Karnopp et al., 2012; Meriam & Kraige, 2012). However, “unpredictable” states are not considered in many of these tools, bringing false models.

Deterministic dynamical mathematical models by themselves are not sufficient to map the system’s randomness because these random variables are exogenous. It boils down to the fact that the perturbation to the state variables is independent (Morrison, 2012). To mitigate this problem, studying the stochastic system has shed light on the robotic system’s field. Last, the class of stochastic systems englobes both continuous and discrete systems. Along with this work, modeling and control are performed in the discrete domain.

**1.1. Deterministic vs. Stochastic System.** Deterministic representation of a model assumes it can predict all the states of a system. However, it is not accurate to practical applications that present uncertain parameters or disturbances (Skogestad & Postlethwaite, 2007). To initiate the discussion of stochastic system control formulation, consider the continuous-time dynamic in the following:

$$(1.1) \quad \dot{x} = f(t, x, u)$$

Equation 1.1 is a mathematical formulation mapping from a set of variables into another, where the control input  $u$  is the action in the system,  $x$  is the states at time  $t$ . Therefore, after excitation of the system with specific control input, and depending on the current states, it will generate a determinist’s response.

Note that the description of this system is deterministic and does not include any uncertainty. Also, it is a continuous-time dynamics. However, it is required for the embedded

\* erivelton.gualter@gmail.com

control system and estimation applications to transform the continuous-time dynamics into a discrete-time dynamic system. Methods such as rectangular, trapezoidal, and Runge-Kutta integration (Simon, 2006) can transform continuous-time to the discrete-time domain, which results in the following representation:

$$(1.2) \quad x_{t+1} = f(t, x_t, u_t), \quad t = 0, 1, \dots, T$$

For a generic system,  $x_t$  is a  $n$ -dimensional state vector and  $u_t$  is a  $m$ -dimensional control vector at time  $t$ .

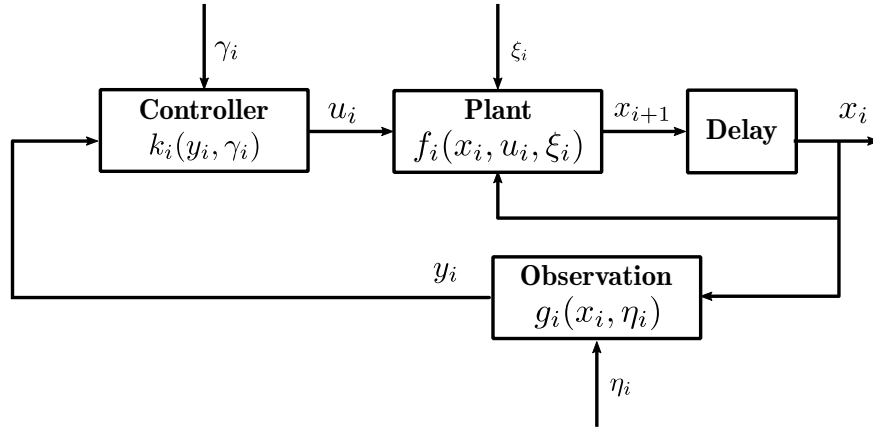
In control engineering, a stochastic system contains random variables which create uncertainties in the modeling. These manifest in terms of the system's parameters, an external noise in the states or control, and random time constant (Aoki, 1967). For instance, the following system contains a random disturbances added to the plant equation:

$$(1.3) \quad x_{i+1} = ax_i + bu_i + \xi_i, \quad u_i \in (-\infty, \infty) \quad 0 \leq i \leq N-1$$

where  $a$  and  $b$  are parameters of the model and  $\xi_i$  is the added disturbances. Generally, the last term are independent with

$$(1.4) \quad \begin{aligned} E(\xi_i) &= 0 \\ E(\xi_i) &= \delta_i^2 \quad 0 \leq i \leq N-1 \end{aligned}$$

The same disturbance can be found on the plant's parameter, control, and states' measurements. Figure 1 illustrates the schematic diagram of the control system. The function form  $f_k$  and  $g_k$  are assumed known; however, the random variables  $\eta_k$ ,  $\gamma_k$ , and  $\xi_k$  are random noise. The Stochastic Control formulation goal is to develop methods to find control input  $k_k$  to drive the states that respect the objective goal. For instance, drive a robot arm on a desired trajectory with less energy consumption, where the goal is to guarantee the robot stays on the path with minimum power effort.



**Figure 1.** Schematic diagram of general stochastic control system.

The formulation of the optimal control used in this work requires a particular stochastic differential equation using **Brownian motion**, also known as **Wiener Process**. According to (Kappen, 2011), this process presents a Gaussian conditional likelihood with a variance that increases linearly in the time interval and has zero mean. Therefore, in order to the following differential equation to be valid, it needs to satisfy  $\langle d\xi d\xi \rangle = v_{ij} dt$ , where  $v$  is a symmetric positive definite matrix.

$$(1.5) \quad dx = f(x(t), u(t), t) dt + d\xi$$

For instance, if the state variable in equation 1.5 is a single state, with means the value of  $v$  is a positive scalar, we have:

$$(1.6) \quad x_{t+1} = x_t + \xi, \quad \xi_t = \pm\sqrt{v}$$

Then, the conditional probability distribution of  $x$  at time  $t$  given initial condition is Gaussian and specified by its mean and variance.

$$(1.7) \quad \rho(x, t | x_0, 0) = \frac{1}{\sqrt{2\pi\nu t}} \exp\left(-\frac{(x - x_0)^2}{2\nu t}\right)$$

## 2. STOCHASTIC OPTIMAL CONTROL

As already stated in the previous section, optimal control formulation's main objective is to drive the states to another state that obeys an objective goal, which can also be found in control engineering literature, as cost function, performance measurement, and others. For instance, if the objective goal is defined to track reference states of a continuous-time system with no perturbation, it can be mathematically described as in the following:

$$(2.1) \quad \begin{aligned} J(x, t) &= \min_u (x - \bar{x})^2 \\ \text{s.t.} \quad \dot{x} &= f(x, u, t) \end{aligned}$$

where  $\bar{x}$  is the desired trajectory of the states. Therefore, the goal is to find an optimal solution to control input that drives the states  $x$  to  $\bar{x}$ . The solution to this optimal problem has been solved by different methods (Kirk, 2004). However, there are uncertainties for the Stochastic system that would result in a slightly different cost function.

The cost function to find an optimal control solution of control input is defined in equation 2.6, which corresponds to the previous stochastic dynamic system's expectation.

$$(2.2) \quad J(t, x) = \mathbb{E}_{\mathbb{Q}} \left[ \phi(x_T, T) + \int_{t_0}^T R(x_t, u_t, t) dt \right]$$

where the *end cost*  $\phi(x_T, T)$  gives the penalty to the last state configuration and integral term is the cost associate to the trajectories  $x(0 \rightarrow T)$ , also known as *optimal cost to go*.

**2.1. Hamilton-Jacobi-Bellman (HJB) solution.** It is necessary to perform a partial derivative (Taylor series expansion) of the cost function with respect to  $dx$  and  $dt$  to obtain the Stochastic Hamilton-Jacobi-Bellman (HJB) equation. Recalling that the stochastic system uses Wiener Process, the partial terms must expand up to order  $dx^2$

$$\begin{aligned}
 \langle J(t+dt, x_{t+dt}) \rangle &= \int dx_{t+dt} \mathcal{N}(x_{t+dt} | x_t, \nu dt) J(t+dt, x_{t+dt}) \\
 (2.3) \quad &= J(t, x_t) + dt \partial_t J(t, x_t) + \langle dx \rangle \partial_x J(t, x_t) + \frac{1}{2} \langle dx^2 \rangle \partial_x^2 J(t, x_t) \\
 &\quad \langle dx \rangle = f(x, u, t) dt \\
 &\quad \langle dx^2 \rangle = \nu(t, x, u) dt
 \end{aligned}$$

Therefore the resulting Stochastic HJB equation is giving in the following:

$$(2.4) \quad -\partial_t J(t, x) = \min_u \left( R(t, x, u) + f(x, u, t) \partial_x J(t, x) + \frac{1}{2} \nu(t, x, u) \partial_x^2 J(t, x) \right)$$

where the boundary condition is  $F(x, T) = \phi(x)$ .

Therefore, the optimal control input is the solution of Equation 2.4. When  $v \mapsto 0$ , the Stochastic HJB reduces to a deterministic HJB problem (Kappen, 2011). The traditional approach for solving linear problems are based on Ricatti equations (Kirk, 2004).

**2.2. Path Integral Control.** The solution of a stochastic optimization problem for a linear system, being discrete or continuous, can be easily solved by Riccati equations (Kirk, 2004). However, a solution for nonlinear is not possible. By transforming the performance measurement (cost function) into an expectation over all possible trajectories, the solution can be obtained with Monte Carlo approximation using forward sampling (Kappen, 2005). The transformation of the cost function is performed with Feynman-Kac lemma, which maps the parabolic partial differential equations (PDEs) to a path integral representation (E. Theodorou et al., 2010).

Before presenting the path integral control solution for stochastic system, the full dynamic system presented in Figure 1 is defined in the following using Brownian disturbance:

$$(2.5) \quad dx = [f(x_t, t) + G(x_t, t) u(x_t, t)] dt + B(x_t, t) dw$$

The cost function also contains a quadratic control cost to minimize the control input, as defined in the following. Note the matrix  $R$  must be a positive definite matrix.

$$(2.6) \quad \mathbf{u}^*(\cdot) = \underset{\mathbf{u}(\cdot)}{\operatorname{argmin}} \mathbb{E}_{\mathbb{Q}} \left[ \phi(x_T, T) + \int_{t_0}^T \mathcal{L}(x_t, \mathbf{u}_t, t) dt \right]$$

Noting that:

$$(2.7) \quad \mathcal{L}(\mathbf{x}_t, \mathbf{u}_t, t) = q(\mathbf{x}_t, t) + \frac{1}{2} \mathbf{u}_t^T \mathbf{R}(\mathbf{x}_t, t) \mathbf{u}_t$$

where the *end cost*  $\phi(\mathbf{x}_T, T)$  gives the penalty to the last state configuration and  $\mathcal{L}(\mathbf{x}_t, \mathbf{u}_t, t)$  is the cost associate to the trajectories  $x(0 \mapsto T)$ , also known as *optimal cost to go*.

In order to estimate how the probability distribution from the uncontrolled  $\mathbb{P}$  and controlled  $\mathbb{Q}$  dynamics, it is used the Girsanov's theorem (Girsanov, 1960). This is called as relative entropy.

$$(2.8) \quad D_{KL}(\mathbb{Q}(\mathbf{u}) \parallel \mathbb{P}) = \frac{1}{2} \int_{t_0}^T \mathbf{u}_t^T \mathbf{G}(\mathbf{x}_t, t)^T \left[ B(\mathbf{x}_t, t) B(\mathbf{x}_t, t)^T \right]^{-1} \mathbf{G}(\mathbf{x}_t, t) \mathbf{u}_t dt$$

Defining  $\Sigma(\mathbf{x}_t, t) = B(\mathbf{x}_t, t) B(\mathbf{x}_t, t)^T$ , and assuming  $\mathbf{R}$  is the product of a scalar positive value  $\lambda$  and  $\mathbf{G}(\mathbf{x}_t, t)^T \Sigma(\mathbf{x}_t, t)^{-1} \mathbf{G}(\mathbf{x}_t, t)$ , we can re-write equation 2.8 as:

$$(2.9) \quad \begin{aligned} D_{KL}(\mathbb{Q}(\mathbf{u}) \parallel \mathbb{P}) &= \frac{1}{2} \int_{t_0}^T \mathbf{u}_t^T \mathbf{G}(\mathbf{x}_t, t)^T \Sigma(\mathbf{x}_t, t)^{-1} \mathbf{G}(\mathbf{x}_t, t) \mathbf{u}_t dt \\ &= \frac{1}{2} \int_{t_0}^T \mathbf{u}_t^T \lambda \mathbf{R} \mathbf{u}_t dt \end{aligned}$$

Therefore, the cost function presented in equation 2.6, can be written as:

$$(2.10) \quad \mathbb{E}_{\mathbb{Q}} \left[ \phi(\mathbf{x}_T, T) + \int_{t_0}^T \mathcal{L}(\mathbf{x}_t, \mathbf{u}_t, t) dt \right] = \mathbb{E}_{\mathbb{Q}} [\phi(\mathbf{x}_T, T) + \lambda D_{KL}(\mathbb{Q}(\mathbf{u}) \parallel \mathbb{P})]$$

The path integral expression presented in the following is the contribution of the relationship between Free Energy and Relative Entropy, which is proof by (E. A. Theodorou, 2015).

$$(2.11) \quad -\lambda \mathcal{F}(S(\tau)) = \inf_{\mathbb{Q}} [\mathbb{E}_{\mathbb{Q}}[S(\tau)] + \lambda D_{KL}(\mathbb{Q} \parallel \mathbb{P})]$$

where  $\mathcal{F}(S(\tau))$  is the free energy distribution over  $\phi(\mathbf{x}_T, T) / \lambda$ :

$$(2.12) \quad \mathcal{F}(S(\tau)) = \log \left( \mathbb{E}_{\mathbb{Q}} \left[ \exp \left( -\frac{1}{\lambda} \phi(\mathbf{x}_T, T) \right) \right] \right)$$

According to E. A. Theodorou (2015), the infimum of the equation 2.11 at differentiation of optimal probability measured of controlled dynamics  $\mathbb{Q}^*$  with respect to probability measure of uncontrolled dynamics  $\mathbb{P}^*$  is:

$$(2.13) \quad \frac{d\mathbb{Q}^*}{d\mathbb{P}} = \frac{\exp\left(-\frac{1}{\lambda}S(\tau)\right)}{\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{1}{\lambda}S(\tau)\right)\right]}$$

Therefore, the computation of the stochastic HJB equation can be avoid by using the relative entropy between  $\mathbb{Q}^*$ , as stated in (Williams et al., 2016):

$$(2.14) \quad \mathbf{u}^*(\cdot) = \operatorname{argmin}_{\mathbf{u}(\cdot)} \mathbb{D}_{KL}(\mathbb{Q}^* \parallel \mathbb{Q}(\mathbf{u}))$$

The formulation of entropy minimization is described in Appendix A and detailed explained in (Williams et al., 2016). Therefore, the resultant entropy is:

$$(2.15) \quad \begin{aligned} \mathcal{D}(\tau, \mathbf{u}(\cdot)) = & - \int_0^T \mathbf{u}_t^T \mathbf{G}(\mathbf{x}_t, t)^T \Sigma(\mathbf{x}_t, t)^{-1} \mathbf{B}(\mathbf{x}_t, t) dw^{(0)} \\ & + \frac{1}{2} \int_0^T \mathbf{u}_t^T \mathbf{G}(\mathbf{x}_t, t)^T \Sigma(\mathbf{x}_t, t)^{-1} \mathbf{G}(\mathbf{x}_t, t) \mathbf{u}_t dt \end{aligned}$$

In conclusion, for a discrete time the optimal control is defined as :

$$(2.16) \quad \mathbf{u}_j^* = \frac{1}{\Delta t} \mathbb{E}_{\mathbb{P}} \left[ \frac{\exp\left(-\frac{1}{\lambda}S(\tau)\right) \mathcal{H}(\mathbf{x}_{t_j}, t_j)}{\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{1}{\lambda}S(\tau)\right)\right]} \right]^{-1} \mathbb{E}_{\mathbb{P}} \left[ \frac{\exp\left(-\frac{1}{\lambda}S(\tau)\right) \mathcal{G}(\mathbf{x}_{t_j}, t_j) \epsilon_j \sqrt{\Delta t}}{\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{1}{\lambda}S(\tau)\right)\right]} \right]$$

where,

$$(2.17) \quad \begin{aligned} \mathcal{G}(\mathbf{x}, t) &= \mathbf{G}(\mathbf{x}, t)^T \Sigma(\mathbf{x}, t)^{-1} \mathbf{B}(\mathbf{x}, t) \\ \mathcal{H}(\mathbf{x}, t) &= \mathbf{G}(\mathbf{x}, t)^T \Sigma(\mathbf{x}, t)^{-1} \mathbf{G}(\mathbf{x}, t) \end{aligned}$$

### 3. SIMULATIONS

The previous section described the stochastic dynamic system's description using Wiener Process and stabilized the optimization control formulation. The critical take-away is the definition of the control law expressed in equations 2.16 and 2.17.

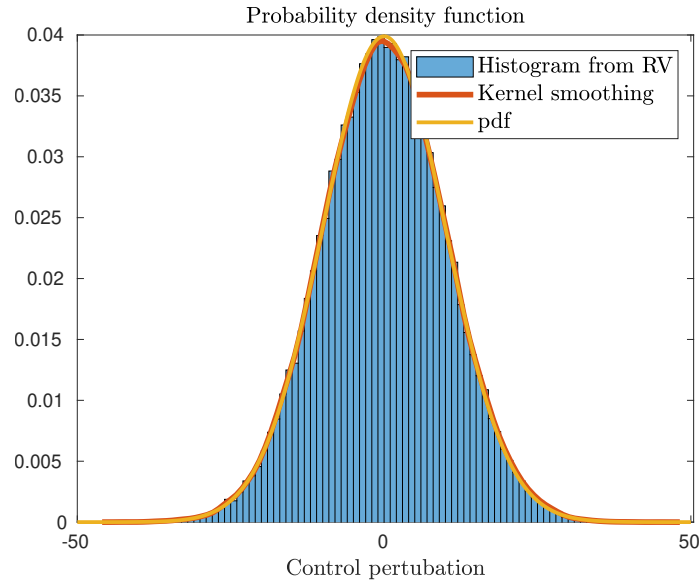
To apply this controller, researchers have been exploring various implementation algorithms, such as Monte-Carlo sampling, to compute the control for the current state (Kapen, 2005) and others. This work takes advantage of an extension of optimal control, called Model predictive control (MPC) (Richter, 2011). MPC is not a specific control strategy. However, it is the artifact to find the control input considering the concept of prediction by minimizing a cost function considering future error and constraints of the problem. See Appendix B for the MPC algorithm used with Path Integral.

The proposed stochastic optimal control is implemented in two system: an inverted pendulum on a cart and a planar two-link robotic arm. The objective goal is to find an optimal control solution to drive the states into a vertical configuration then balance.

Both applications contain perturbation  $\gamma_i$  in the control input as:

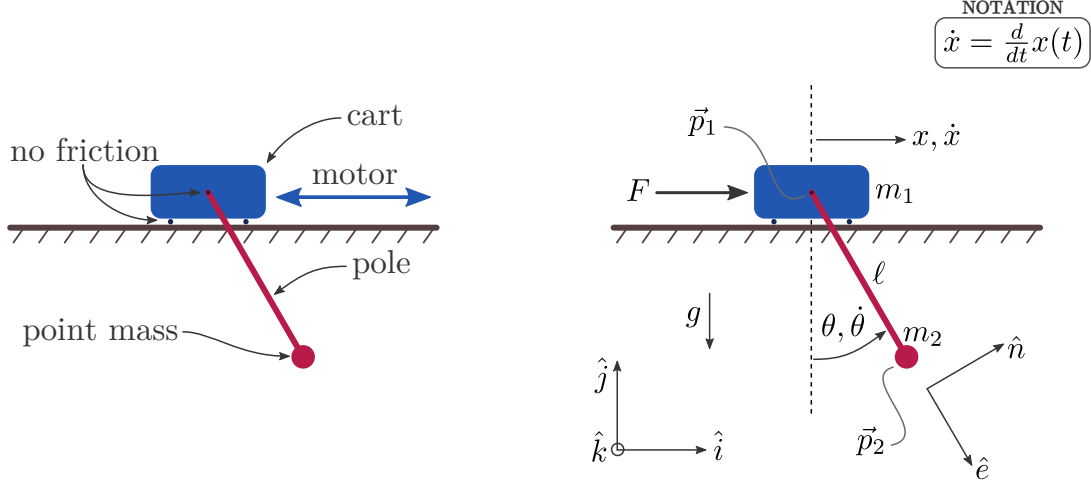
$$(3.1) \quad \begin{aligned} E(\gamma_i) &= 0 \\ E(\gamma_i) &= \delta_i^2 \quad 0 \leq i \leq N-1 \end{aligned}$$

Figure 2 shows the probability density function for the system's random uncertainty. The histogram contains the actual perturbation in the system. Kernel smoothing plot is an estimation of the pdf given the random variables. Last, the pdf plot (in yellow) corresponds to the pdf with a provided variance and mean.



**Figure 2.** Probability density function.

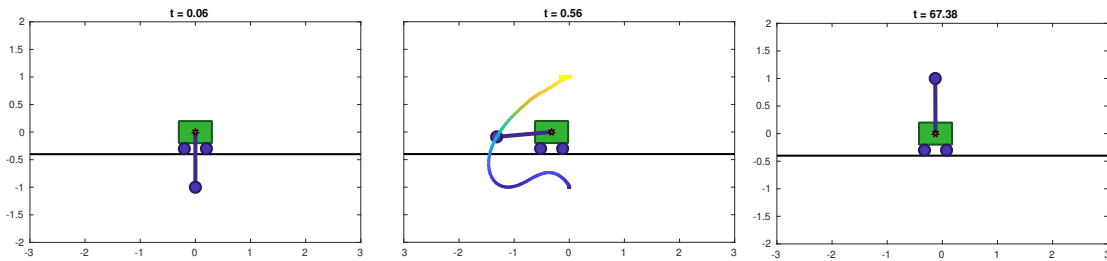
**3.1. Inverted Pendulum on a cart.** A well-known problem in control theory is presented in the following to illustrate the contraction and introduce the stochastic optimal control solution of this work: the inverted pendulum problem. Figure 7 consists of a motorized cart and a bar. The mathematical description of the problem is analog to technologies found on a rocket launch (Ogata, 2010) and two-wheeled self-balancing personal transporter (Chan et al., 2013). The governing equation is represented in 3.2:

**Figure 3.** Inverted Pendulum System.

The control input  $u$  is the force  $F$  that moves the cart, and the outputs are the angular position  $\theta$  and the horizontal position of the cart  $x$ . For the optimal control formulation presented in this work, the author uses a state-space presentation, where the states are  $x = [x \ \dot{x} \ \theta \ \dot{\theta}]$ . In control engineering, it corresponds to a system's mathematical model as a set of input, output, and state variables related by first-order differential equations.

$$(3.2) \quad \begin{pmatrix} \cos \theta & \ell \\ m_1 + m_2 & m_2 \ell \cos \theta \end{pmatrix} \begin{pmatrix} \ddot{x} \\ \ddot{\theta} \end{pmatrix} = \begin{pmatrix} -g \sin \theta \\ F + m_2 \ell \dot{\theta}^2 \sin \theta \end{pmatrix}$$

The results of stochastic optimization control is presented in Figures 4, 5, and 6. The plot on the left shows the initial state. The plot on the right shows the final state. The plot in between shows an intermediate step to achieve the objective control. The curve presents the path of the endpoint. Note that the control input is only the trust force of the card.

**Figure 4.** Optimal trajectory for the Inverted Pendulum System on a cart.



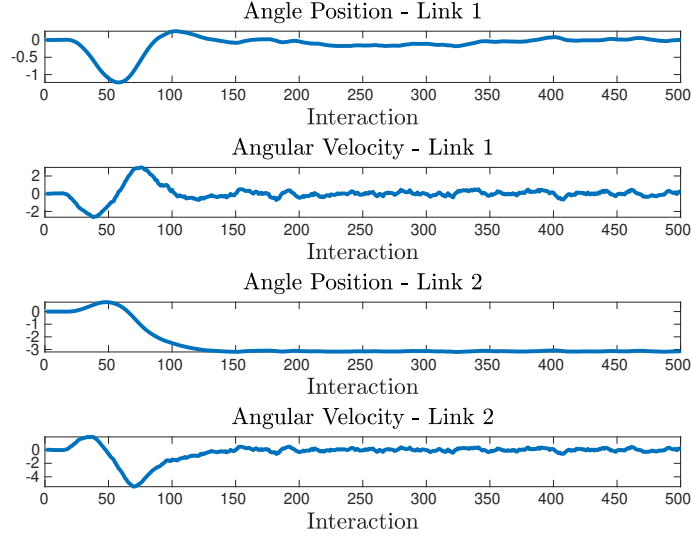


Figure 5. State Space Results

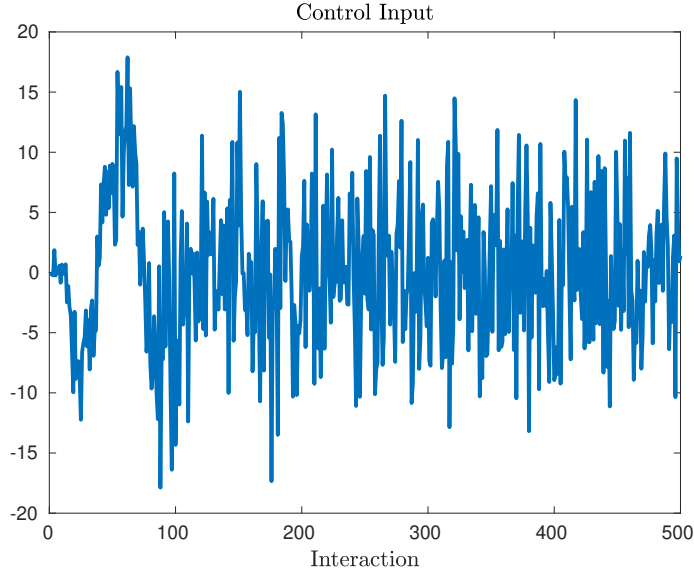
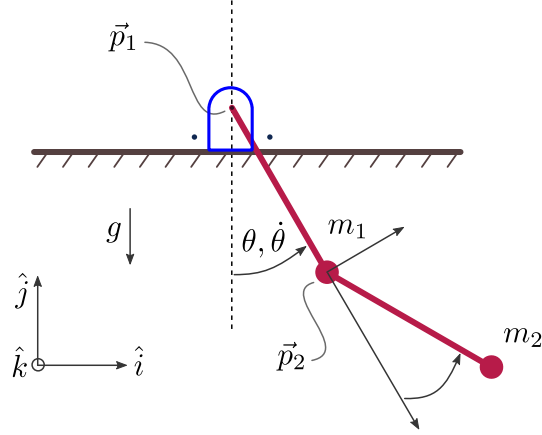


Figure 6. Optimal Control solutions.

**3.2. Planar two-link robotic arm.** A planar two-link robotic arm, also known as an acrobot, is composed of two rigid links connected by a two-degree rotational joint. In the control engineering literature, the common control task studied for this robot configuration is the swing-up task, aiming to find an optimal control solution to move the system into a vertical configuration and then balance.



**Figure 7.** Planar two-link robotic arm.

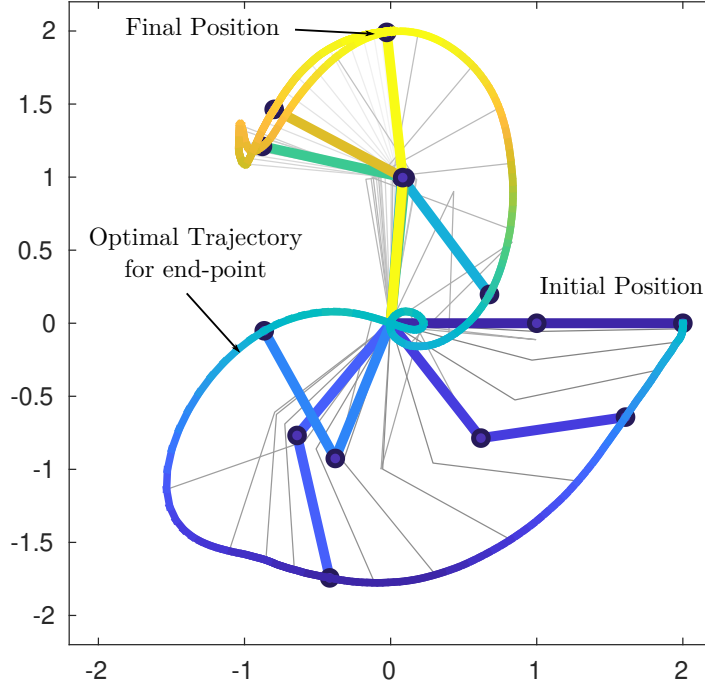
The governing equation of a two-link robotic arm might be represented in the standard robotic formulation Tedrake, 2016:

$$(3.3) \quad \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} = \tau_g(\mathbf{q}) + \mathbf{B}\mathbf{u}$$

where:

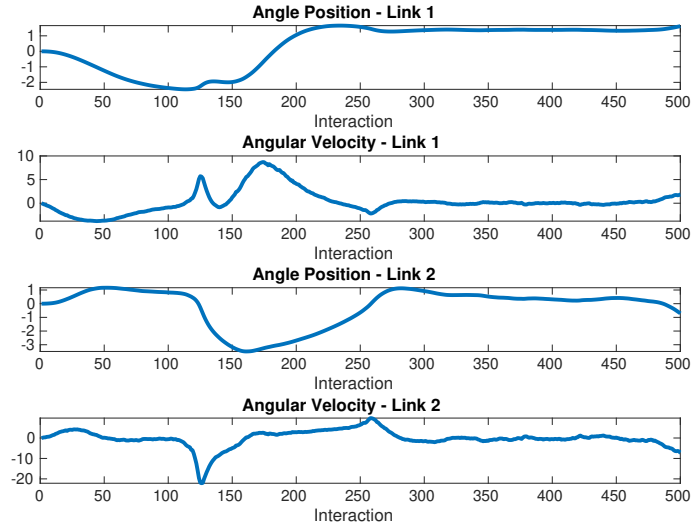
$$(3.4) \quad \begin{aligned} \mathbf{M}(\mathbf{q}) &= \begin{bmatrix} I_1 + I_2 + m_2 l_1^2 + 2m_2 l_1 l_{c2} c_2 & I_2 + m_2 l_1 l_{c2} c_2 \\ I_2 + m_2 l_1 l_{c2} c_2 & I_2 \end{bmatrix} \\ \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) &= \begin{bmatrix} -2m_2 l_1 l_{c2} s_2 \dot{q}_2 & -m_2 l_1 l_{c2} s_2 \dot{q}_2 \\ m_2 l_1 l_{c2} s_2 \dot{q}_1 & 0 \end{bmatrix} \\ \tau_g(\mathbf{q}) &= \begin{bmatrix} -m_1 g l_{c1} s_1 - m_2 g (l_1 s_1 + l_{c2} s_{1+2}) \\ -m_2 g l_{c2} s_{1+2} \end{bmatrix} \\ \mathbf{B} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

Note, the presented governing equation is a continuous-time nonlinear dynamical system. Therefore, for this project's conclusion, a discretization of the system is performed, and random perturbation is added to control and parameters.

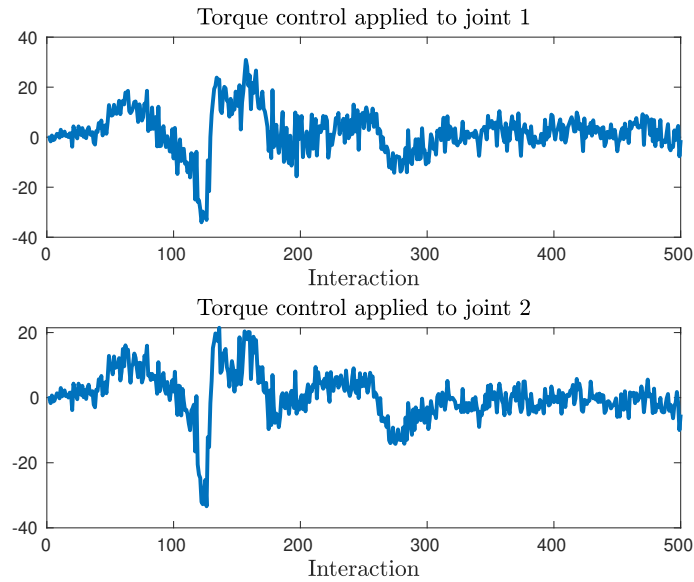


**Figure 8.** Optimal trajectory for the planar two-link robotic arm.

The optimal solution is presented in Figures 8, 9, and 10. Figure 8 shows some snapshots during animation of the swing-up maneuver. Note, the robot starts with the joints to the right, and it moves to the left to storage potential energy and moves up.



**Figure 9.** State Space Results



**Figure 10.** Optimal Control solutions.

#### 4. MATLAB IMPLEMENTATION

All the code is available at: <https://github.com/EriveltonGualter/ESC-704-Stochastic-Systems>.

To compute the entered problem, run the MAIN.m file:

```
1 run('MAIN.m')
```

Otherwise, if it is desired only show the plots, run the MAINPresentation.m file:

```
1 run('MAIN_Presentation.m')
```

## APPENDIX A. RELATIVE ENTROPY MINIMIZATION

The optimal control solution using path integral method has been greatly presented in Williams et al. (2016) and here is just described the high level derivations.

First, it is found the relative entropy between the optimal distribution  $Q^*$  and induced distribution  $Q(u)$  of the control:

$$(A.1) \quad \mathbb{D}_{KL}(Q^* \| Q(\mathbf{u})) = \mathbb{E}_{Q^*} \left[ \log \left( \frac{dQ^*}{dQ(\mathbf{u})} \right) \right]$$

After applying the chain rule property of Radon-Nikodym derivatives and Girsanov's theorem, we have:

where  $\frac{d\mathbb{P}}{dQ(\mathbf{u})}$  is defined as:

$$(A.2) \quad \frac{d\mathbb{P}}{dQ(\mathbf{u})} = \exp(\mathcal{D}(\tau, \mathbf{u}(\cdot)))$$

and

$$(A.3) \quad \begin{aligned} \mathcal{D}(\tau, \mathbf{u}(\cdot)) = & - \int_0^T \mathbf{u}_t^T \mathbf{G}(\mathbf{x}_t, t)^T \Sigma(\mathbf{x}_t, t)^{-1} \mathbf{B}(\mathbf{x}_t, t) dw^{(0)} \\ & + \frac{1}{2} \int_0^T \mathbf{u}_t^T \mathbf{G}(\mathbf{x}_t, t)^T \Sigma(\mathbf{x}_t, t)^{-1} \mathbf{G}(\mathbf{x}_t, t) \mathbf{u}_t dt \end{aligned}$$

Combining with equation 2.13, the entropy between the optimal distribution  $Q^*$  and probability distribution from the uncontrolled dynamics  $P$  is defined as:

$$(A.4) \quad \mathbb{D}_{KL}(Q^* \| Q(\mathbf{u})) = \mathbb{E}_{Q^*} \left[ \log \left( \frac{\exp \left( -\frac{1}{\lambda} S(\tau) \right) \exp(\mathcal{D}(\tau, \mathbf{u}(\cdot)))}{\mathbb{E}_{\mathbb{P}} \left[ \exp \left( -\frac{1}{\lambda} S(\tau) \right) \right]} \right) \right]$$

From section 2.2, the free energy distribution over  $\phi(\mathbf{x}_T, T)/\lambda$ , which correspond to the *end cost*, does not depend of control. Therefore, the terms regarding this cost might be neglected. The minimization problem results in:

$$(A.5) \quad \underset{\mathbf{u}(\cdot)}{\operatorname{argmin}} \mathbb{D}_{KL}(Q^* \| Q(\mathbf{u})) = \underset{\mathbf{u}(\cdot)}{\operatorname{argmin}} \mathbb{E}_{Q^*}[\mathcal{D}(\tau, \mathbf{u}(\cdot))]$$

## APPENDIX B. MODEL PREDICTIVE PATH INTEGRAL CONTROL ALGORITHM

The MPC control works differently from the conventional controllers. Instead of controlling the plant-based in trying to minimize the past error, it simulates a series of candidate control inputs in the short future, which minimizes the desired cost function. When combining with scholastic optimal control formulation, the algorithm can be described in Figure 11 (Williams et al., 2016).

**Algorithm 1:** Model Predictive Path Integral Control

---

**Given:**  $K$ : Number of samples;  
 $N$ : Number of timesteps;  
 $(\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{N-1})$ : Initial control sequence;  
 $\Delta t, \mathbf{x}_{t_0}, \mathbf{f}, \mathbf{G}, \mathbf{B}, \mathbf{B}_E$ : System/sampling dynamics;  
 $\phi, q, \mathbf{R}, \lambda$ : Cost parameters;  
 $\mathbf{u}_{\text{init}}$ : Value to initialize new controls to;

**while** *task not completed* **do**

**for**  $k \leftarrow 0$  **to**  $K - 1$  **do**

$\mathbf{x} = \mathbf{x}_{t_0}$ ;

**for**  $i \leftarrow 1$  **to**  $N - 1$  **do**

$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{f} + \mathbf{G}\mathbf{u}_i)\Delta t + \mathbf{B}_E\epsilon_{i,k}\sqrt{\Delta t}$ ;

$\tilde{S}(\tau_k) = \tilde{S}(\tau_k) + \tilde{q}(\mathbf{x}_i, \mathbf{u}_i, \epsilon_{i,k}, t_i)$ ;

**for**  $i \leftarrow 0$  **to**  $N - 1$  **do**

$\mathbf{u}_i =$

$\mathbf{u}_i + \mathcal{H}^{-1} \mathcal{G} \left[ \sum_{k=1}^K \left( \frac{\exp(-\frac{1}{\lambda} \tilde{S}(\tau_k)) \frac{\epsilon_{j,k}}{\sqrt{\Delta t}}}{\sum_{k=1}^K \exp(-\frac{1}{\lambda} \tilde{S}(\tau_k))} \right) \right]$ ;

send to actuators( $\mathbf{u}_0$ );

**for**  $i \leftarrow 0$  **to**  $N - 2$  **do**

$\mathbf{u}_i = \mathbf{u}_{i+1}$ ;

$\mathbf{u}_{N-1} = \mathbf{u}_{\text{init}}$

Update the current state after receiving feedback;

check for task completion;

---

Figure 11. MPC-Path Integral.

## REFERENCES

- Aoki, M. (1967). *Optimization of stochastic systems: Topics in discrete-time systems*. Academic Press.
- Chan, R. P. M., Stol, K. A., & Halkyard, C. R. (2013). Review of modelling and control of two-wheeled robots. *Annual reviews in control*, 37(1), 89–103.
- Girsanov, I. V. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, 5(3), 285–301.
- Kappen, H. J. (2005). Path integrals and symmetry breaking for optimal control theory. *Journal of statistical mechanics: theory and experiment*, 2005(11), P11011.
- Kappen, H. J. (2011). Optimal control theory and the linear bellman equation.
- Karnopp, D. C., Margolis, D. L., & Rosenberg, R. C. (2012). *System dynamics: Modeling, simulation, and control of mechatronic systems*. John Wiley & Sons.
- Kirk, D. E. (2004). *Optimal control theory: An introduction*. Courier Corporation.
- Meriam, J. L., & Kraige, L. G. (2012). *Engineering mechanics: Dynamics* (Vol. 2). John Wiley & Sons.
- Morrison, F. (2012). *The art of modeling dynamic systems: Forecasting for chaos, randomness and determinism*. Courier Corporation.
- Ogata, K. (2010). *Modern control engineering*. Prentice hall.
- Richter, H. (2011). *Advanced control of turbofan engines*. Springer Science & Business Media.
- Simon, D. (2006). *Optimal state estimation: Kalman, h infinity, and nonlinear approaches*. John Wiley & Sons.
- Skogestad, S., & Postlethwaite, I. (2007). *Multivariable feedback control: Analysis and design* (Vol. 2). Citeseer.
- Tedrake, R. (2016). Underactuated robotics: Algorithms for walking, running, swimming, flying, and manipulation. *Course Notes for MIT*, 6.
- Theodorou, E., Buchli, J., & Schaal, S. (2010). A generalized path integral control approach to reinforcement learning. *The Journal of Machine Learning Research*, 11, 3137–3181.
- Theodorou, E. A. (2015). Nonlinear stochastic control and information theoretic dualities: Connections, interdependencies and thermodynamic interpretations. *Entropy*, 17(5), 3352–3375.
- Williams, G., Drews, P., Goldfain, B., Rehg, J. M., & Theodorou, E. A. (2016). Aggressive driving with model predictive path integral control. *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 1433–1440.