

# Socially Aware Motion Planning for a Flying Robot with Model Predictive Path Integral Control

Hyung-Jin Yoon, Pan Zhao, Chuyuan Tao, Christopher Widdowson,  
Ranxiao F. Wang, Naira Hovakimyan and Evangelos A. Theodorou

**Abstract**—This paper presents a motion control framework for a flying robot that takes into account the safety perception of humans in close proximity. Human’s safety perception is predicted based on data collected from physiological experiments in a virtual reality (VR) environment. The predicted safety perception is incorporated in optimal control of the robot’s motion based on model predictive path integral (MPPI) control. Compared to our previous work [1] that addressed off-line path planning for a flying robot incorporating perceived safety, the MPPI-based framework in this paper is able to generate perceived-safe motion of the robot online, accounting for changes in the environments, e.g. human position change, in real time. This makes it applicable to dynamically changing environments. The proposed framework is demonstrated in a VR environment with a human in the loop.

## I. INTRODUCTION

In the recent decade, multi-rotor copters have gone outside the lab space and seen immense growth in the commercial market for various civilian applications such as media production, inspection, and precision agriculture. More ambitious applications, e.g. for package delivery, fire fighting, elderly care are also under investigation [2]–[4]. The increasing popularity of these flying robots can be attributed to their mechanical simplicity, ability to hover and high maneuverability. It is expected that by 2020, the market for these devices will attain a value of \$11.2 billion with an annual growth of over 30% [5]. The inclusion of these aerial robots in our day-to-day lives brings immediate benefits to society as well as individuals. However, since they often need to fly near people and navigate in densely populated areas, these robots also pose significant challenges in terms of ensuring human safety and comfort during task execution.

It is a long tradition in robot control and motion planning to focus on the robot’s safety, i.e., the ability to generate safe collision-free paths. However, this is insufficient for robots operating in proximity to humans, who can still feel unsafe or uncomfortable even though the robot is guaranteed not to collide with them. Studies of human perception have shown that there is a sharp distinction between human perceived safety and actual safety.

Research supported by NSF NRI initiative #1528036 and #1830639.

Hyung-Jin Yoon, Pan Zhao, Chuyuan Tao, and Naira Hovakimyan are with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign (UIUC), Urbana, IL 61801, USA. Christopher Widdowson and Ranxiao Wang are with Psychology Department and Beckman Institute for Advanced Science and Technology in UIUC. Evangelos Theodorou is with Aerospace Engineering of Georgia Tech; {hyoon33, panzhao2, chuyuan2, widdwsn2, wang18, nhovakim}@illinois.edu, evangelos.theodorou@gatech.edu

A comfortable approach distance for a flying robot was studied in [6] using a real robot, where the authors found no conclusive effects of the height of the robot as well as the comfortable distance. In [7], the influence of the robot’s speed and repeating behavior (cyclicity) on distancing with the robot and the interaction preference of humans was investigated using virtual reality (VR) experiments.

On the other hand, a few design approaches to improving humans’ comfort and acceptance of these flying robots have been explored. Emotional encoding was proposed in [8], where the authors suggested adding emotional components to the drones to facilitate natural human-drone interaction. In [9], the authors developed a set of signaling mechanisms for a flying robot to better communicate its intent (i.e. directionality, which were tested on real robots and human participants). Also in our recent work [1], we presented a path planning framework that considers the human’s safety perception for flying robots and the comfort of the users when they interact with flying robots.

The work mentioned above focuses on either qualitatively studying the influence of various features of the robot on human-aerial-robot interaction, or devising heuristic methods to improve the acceptability of the robot by humans. In our recent work [1], we first studied and modeled the dependence of human’s perceived safety on an aerial robot’s position and velocity by measuring the arousal signal of humans in response to a flying robot in a VR environment. We then developed a path planning framework that incorporates the human’s safety perception model and generates paths that are felt safe and comfortable by the humans.

This paper is an extension of our previous work [1]. The contributions are listed as follows:

- 1) Our previous work [1] can only do off-line trajectory generation due to the computationally-expensive nonlinear optimization in the optimal control problem formulation, which is not suitable for dynamically varying environments. We extend it for online trajectory generation by invoking model predictive path-integral (MPPI) control [10].
- 2) The proposed control method is implemented and tested in a VR environment with a human in the loop to demonstrate the socially aware behavior of the robot, while the path planning framework in our previous work [1] is only implemented in a simulation.

The remainder of the paper is organized as follows. Section II introduces the data-driven model of the human’s safety perception based on the data from VR experiments [1].

In Section III, we present the main framework, in which the safety perception model is incorporated into the MPPI control. In Section IV, the proposed framework is implemented and tested in the VR environment. Section V summarizes and discusses future directions.

## II. DATA-DRIVEN MODELLING OF HUMANS' SAFETY PERCEPTION

In this section, we briefly review the data-driven modeling of humans' safety perception, which was already presented in [1].

### A. Virtual Reality based Data Collection

VR offers a safe, low-cost, and time efficient method to collect data [7]. To this end, we have developed a VR test environment to explore human-aerial-robot interactions in a variety of experimental scenarios [11], [12]. Concurrent psychophysiological reactions of participants are recorded in terms of head motions and electrodermal activity (EDA), and time-aligned with attributes of the robot's flight path, e.g. velocity, altitude, and audio profiles. Participants were seated at the junction of a three-way intersection with unoccluded paths in the forward, left, and right direction (see Fig. 1). Three arbitrary trajectories conforming to the shape of the intersection were chosen and reversed, for a total of six unique trajectories (1.6 m altitude). We collected the data from 56 participants (20 males / 36 females).



Fig. 1: A flying robot in the VR environment (an illustration video at <https://youtu.be/XnaXzdH1xUA>).

We measured the EDA using skin conductance sensor, the data from which is preprocessed by EDA analysis package, Ledalab, to generate the phasic activation signal [13]. As argued in [1], although physiological measurements of arousal (e.g., EDA) alone are not necessarily equivalent to perceived safety, several pieces of evidence suggest that the EDA measure of physiological response in our study is closely related to people's anxiety induced by the approaching drone. For all experiments, the EDA signal was stronger when the quadrotor was flying at a higher speed, with the audio on, and at eye-height rather than overhead, but this trend decreased across subsequent trials. Also, head motion accelerated in the direction opposite of the quadrotor's motion on its approach, indicating avoidance behavior. Thus, in the following sections, we consider the EDA signal as an operational approximation of the human's perceived safety.

### B. The Safety Perception Model

A data-driven model was also developed in [1] to predict the phasic activation (arousal), given the robot's position and velocity. Let  $y_n \in \mathbb{R}$  denote the phasic activation, where  $n$  is the time index. The input (feature) variable, denoted by

$x_n \in \mathbb{R}^8$ , is the vector that contains the distance to the robot, the rate of change of the distance, the Cartesian position coordinates, and the velocity coordinates of the robot relative to the human.

It was found that the phasic activation of a subject was often influenced by some unknown factors that are not related to the drone. Figure 2 shows the presence of such unknown factors in one data set, where an increase in the phasic activation in the shaded area was detected, although the flying robot was far away and virtually invisible to the subject. To account for the unknown factors in the data, we

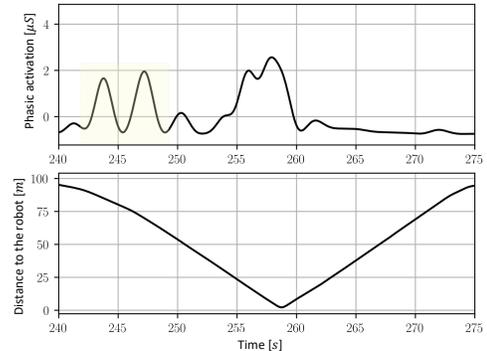


Fig. 2: Phasic activation signal induced by the flying robot. The shaded box indicates the response, where the robot is in the far distance (greater than 60 [m]).

hypothesize that the unexpected spike of the phasic activation in Figure 2 is due to the change of the participant's attention, i.e. the participant is distracted by some other stimulus. Inspired by the work in [14], we model the sequential dependence of the (hidden) human's attention using a Hidden Markov Model (HMM). The HMM has a latent variable<sup>1</sup>  $Z_n$  to denote two states of human's attention:

$$Z_n := \begin{cases} 1, & \text{if the human is attentive to the robot,} \\ 2, & \text{otherwise.} \end{cases}$$

Then  $Z_n$  is modeled by a Markov chain with the following probability transition equation:

$$\pi_{n+1} = \pi_n \mathbf{A}. \quad (1)$$

The vector  $\pi_n := [\mathbb{P}(Z_n = 1), \mathbb{P}(Z_n = 2)]$  is the stochastic row vector for the distribution over the state  $Z_n$ , and  $\mathbf{A} \in \mathbb{R}^{2 \times 2}$  denotes the transition probability matrix of the Markov chain. The initial distributions  $\pi_0$  and  $\mathbf{A}$  are the parameters of the Markov chain. The output is determined as

$$Y_n = \mathbb{1}_{\{Z_n=1\}}(f_\beta(X_n) + \varepsilon) + \mathbb{1}_{\{Z_n=2\}}\delta, \quad (2)$$

where  $\mathbb{1}_A$  denotes the indicator function, and  $f_\beta: \mathbb{R}^8 \rightarrow \mathbb{R}$  is a function  $f_\beta(x) := \beta^\top \phi(x)$ , which is linearly parameterized with  $\beta$  and basis  $\phi(x)$ ,  $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ , and  $\delta$  denotes the random source. As seen in (2),  $Y_n$  depends on  $X_n$  when  $Z_n = 1$ ; however,  $Y_n = \delta$  when  $Z_n = 2$ , i.e. it is modeled as an

<sup>1</sup>In this paper, we interchangeably use capital letters to denote random variables and lower case letter to denote deterministic variables (or samples).

independent random signal. In (2), it can be seen that one of the two regression functions,  $Y_n = f_\beta(X_n)$  and  $Y_n = \delta$ , is chosen based on the likelihood given the observation  $(x_n, y_n)$ .

The parameter of the model was estimated by the maximum likelihood estimation (MLE) algorithm using Expectation-Maximization (EM) [15] and was given in [1].

### III. MODEL PREDICTIVE MOTION CONTROL WITH PERCEIVED SAFETY

The trajectory generation method in [1] is not suitable for dynamically changing environments. For example, if the position of the human changes, then the optimal trajectory calculated based on the method in [1] will not be applicable anymore. Hence, a feedback controller which takes into account the current state is desired in such environments. Model predictive control (MPC) has been widely used for control tasks that involve minimization of a cost function for performance optimization while satisfying some constraints [16]. In MPC, a sequence of optimal control inputs,  $u_t, u_{t+1}, \dots$ , is solved for a finite time horizon at every time step  $t$  accounting for the current state  $x_t$ , while usually only the first element of the control input sequence, i.e.  $u_t$ , is executed. At the next time step  $t+1$ , the above procedure is repeated. This receding-horizon feature of MPC introduces feedback and makes it capable of handling dynamically varying environments.

The safety perception model presented in Section II can be used to create a cost function (or a constraint) in MPC, which penalizes the robot's paths that induce discomfort to the human. MPC is a control technique that has been widely used for accommodating different constraints. For the motion control incorporating perceived safety, we choose the model predictive path-integral (MPPI) control.

#### A. Brief Introduction to MPPI

The MPPI control algorithm solves stochastic optimal control problem based on the (stochastic) sampling of the system trajectories through parallel computation [10], [17], [18]. Due to the sampling nature, the algorithm does not require derivatives of either the dynamics or the cost function of the system, which enables to handle nonlinear dynamics and non-smooth/non-differentiable cost functions without approximations. With the help of GPUs for expediting the parallel computation, the MPPI can be implemented in real time even for relatively large dimensions of the state space (e.g. there are 48 state variables for the 3-quadrotor control example in [10]). The computational efficiency from paralleled stochastic sampling and the ability to directly handle non-smooth cost functions (cf. equation (5)) make MPPI appealing for the motion control problem in this paper.

#### B. MPPI Motion Control with Perceived Safety

Note that MPPI control can handle complex and/or nonlinear dynamics. However, to facilitate the comparison with our previous work [1] that uses a simple double integrator model for the robot, in pure (i.e. not VR) simulation the double integrator model is used again here. In the VR demonstration

---

#### Algorithm 1 Model Predictive Path Integral [17]

---

##### Given parameters:

$F$ : Transition model given in (4);

$S$ : Cost function given in (6);

$T$ : Length of the finite horizon;

##### Choose tuning parameters:

$K$ : Number of sample trajectories;

$\Sigma$ : Co-variance of the noise  $\varepsilon_k$ ;

$\lambda$ : Temperature parameter of the Gibbs distribution in (3);

$\bar{\mathbf{u}} := (\mathbf{u}_1, \dots, \mathbf{u}_T)$ : Initial control sequence;

$\mathbf{u}_{\text{new}}$ : Value to initialize new control to;

##### repeat

0. Measure current state and save as  $\mathbf{x}_{\text{init}}$ .

1. Sample  $K$  trajectories of noise,

$$\bar{\varepsilon}^{(k)} := (\varepsilon_1^{(k)}, \dots, \varepsilon_T^{(k)}), \quad \varepsilon_t^{(k)} \sim \mathcal{N}(0, \Sigma).$$

2. Roll out  $K$  sample trajectories with  $\bar{\varepsilon}^{(k)}$ ,  $\bar{\mathbf{u}}$ , and  $\mathbf{x}_0$  and calculate the cost  $S^{(k)}$ .

$$\mathbf{x}_{t+1}^{(k)} = \mathbf{A}\mathbf{x}_t^{(k)} + \mathbf{B}(\mathbf{u}_t + \varepsilon_t^{(k)}), \quad \mathbf{x}_1^{(k)} = \mathbf{x}_{\text{init}}.$$

3. Calculate estimated optimal control using the costs calculated from the  $K$  trajectories,  $(S^{(1)}, S^{(2)}, \dots, S^{(K)})$ , as in (6) and

$$\mathbf{u}_t \leftarrow \mathbf{u}_t + \sum_{k=1}^K \omega_k \varepsilon_t^{(k)}, \quad \forall t \in \{1, 2, \dots, T\},$$

where

$$\omega_k = \frac{\exp(-(S^{(k)} - \rho)/\lambda)}{\sum_{k=1}^K \exp(-(S^{(k)} - \rho)/\lambda)}, \quad \rho := \min_k S^{(k)}. \quad (3)$$

4. Send  $\mathbf{u}_1$  to the actuator.

5. Update the control sequence  $\bar{\mathbf{u}}$  as

$$\bar{\mathbf{u}} \leftarrow (\mathbf{u}_2, \dots, \mathbf{u}_{T-1}, \mathbf{u}_{\text{new}}).$$

**until** mission ends.

---

we use a different model, obtained from system identification of the dynamics of the robot in the VR, as explained in Section IV. The discrete-time double integrator model for the simplified planner dynamics of the flying robot is defined as

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}(\mathbf{u}_t + \varepsilon_t), \quad (4)$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix},$$

where  $\mathbf{x}_t \in \mathbb{R}^4$  denotes the state vector consisting of planar position and velocity coordinates,  $\mathbf{u}_t \in \mathbb{R}^2$  denotes the control input which is planar acceleration,  $\varepsilon_t \in \mathbb{R}^2$  is independent and identically distributed (i.i.d.) noise with normal distribution, i.e.  $\varepsilon_t \sim \mathcal{N}(0, \Sigma)$  with the co-variance matrix  $\Sigma$ , and  $\Delta t$  is the sampling interval, which is set to 0.1 second in this section.

We select the running cost as

$$c(\mathbf{x}) = 1000(\mathbb{1}_{C_1}(\mathbf{x})) + 1000(\mathbb{1}_{C_2}(\mathbf{x})), \quad (5)$$

where  $\mathbb{1}_C: X \rightarrow \{0, 1\}$  is the indicator function, and

$$C_1 = \{\mathbf{x} | f_\beta(\mathbf{x}) \geq b_a\}, \quad C_2 = \{\mathbf{x} | \text{collision with an obstacle}\}.$$

The first term in (5) is for penalizing the states that will make a human feel unsafe, represented by the predicted phasic activation exceeding some threshold  $b_a$ , calculated with  $f_\beta(\cdot)$  in (2). The second term is for avoiding collisions with obstacles. The terminal cost is selected to encourage reaching the goal with a minimum velocity, as

$$\phi(\mathbf{x}) = (x - x_{\text{goal}})^2 + (y - y_{\text{goal}})^2 + v_x^2 + v_y^2,$$

where  $\mathbf{x} = [x, y, v_x, v_y]^\top$ , and  $(x_{\text{goal}}, y_{\text{goal}})$  is the target position. Given a trajectory of state  $\bar{\mathbf{x}} := (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T)$  and control  $\bar{\mathbf{u}} := (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_T)$ , the cost of the trajectory,  $S(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ , is defined as

$$S(\bar{\mathbf{x}}, \bar{\mathbf{u}}) = \phi(\mathbf{x}_T) + \sum_{t=1}^T \left( c(\mathbf{x}_t) + \lambda \mathbf{u}_k^\top \Sigma^{-1} \mathbf{e}_t \right), \quad (6)$$

where  $\lambda$  is the temperature parameter of the Gibbs distribution (or Softmax function) used in (3). Algorithm 1 describes how the MPPI calculates the control input at every time step [17]. Figure 3 shows the resultant flight paths based on MPPI. Similar to [1], the threshold  $b_a$  can be used as a tuning knob to adjust the allowable discomfort level induced by the robot to a nearby human as shown in Figure 3.

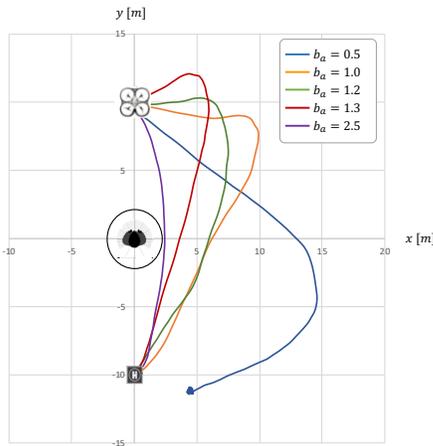


Fig. 3: Flight paths for different threshold  $b_a$ . The chosen MPPI parameters are as follows:  $K = 2000, \Sigma = I, \lambda = 1, T = 50$ , and  $\bar{\mathbf{u}}$  is set to zero vectors.

#### IV. VIRTUAL REALITY DEMONSTRATION

In this section, we implemented the MPPI based framework in a VR environment to validate its use for the socially aware motion control of a flying robot. The Unity 3D [19], a game development editor, was used to construct the VR environment, where the physics engine of the editor was able to conveniently model the rigid-body dynamics of the

flying robot. We ran system identification to determine  $\mathbf{A}$  and  $\mathbf{B}$  as the MPPI control presented in III needs the model parameters. The sampling frequency for implementing the MPPI control was selected as 25 Hz. An illustration of the test environment is shown in Figure 4 and a video demonstration of the results are available at <https://youtu.be/UtWrSTDAZsw>. We compared the pure motion control with the collision avoidance using the MPPI framework (by removing the perceived-safety-related term from the cost function (5)) with the socially aware control incorporating the safety perception (with  $b_a = 1.0$ ). The comparison results are included in the video, from which one can see that the flight path from the socially aware control keeps a larger distance from the human by following a smooth arc, as opposed to the straight-line path yielded by the collision-avoidance motion control. Furthermore, as also demonstrated in the video, the robot was able to react to the varying position of the human in the VR and re-plan its path, thanks to the feedback nature of the MPPI framework.

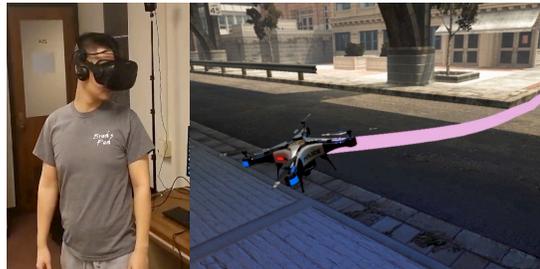


Fig. 4: A flying robot under the MPPI control in the VR environment. The colored trailing line is to visualize the path.

#### V. CONCLUSION

We presented a framework for optimal motion control of a flying robot incorporating the safety perception of humans in its neighborhood. The framework builds upon model predictive path integral (MPPI) control [10] for online optimal control and human's safety perception model. It extends the earlier results from [1] for off-line trajectory generation with the account of human's safety perception to online setting. The capacity of the proposed framework to generate socially-aware paths that try to avoid human discomfort as well as reacting to the change of environments was demonstrated in virtual-reality based experiments with the human in the loop.

Same as our previous work [1], the proposed framework employs a fixed safety perception model, which may not well account for the differences between humans and the variation of a human's safety perception. Our future work will be investigating adaptive optimal control framework for socially-aware motion planning and control for the flying robot in the vicinity of humans, by invoking the hidden Markov model-based reinforcement learning in [20].

#### APPENDIX

The accompanying video is available at <https://youtu.be/dDAfr9qkmmg>.

## REFERENCES

- [1] H.-J. Yoon, C. Widdowson, T. Marinho, R. F. Wang, and N. Hovakimyan, "A path planning framework for a flying robot in close proximity of humans," *arXiv preprint arXiv:1903.05156*, 2019.
- [2] R. D'Andrea, "Guest editorial can drones deliver?" *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 3, pp. 647–648, 2014.
- [3] J. Cohen, "Drone spy plane helps fight california fires," 2007.
- [4] T. Marinho, C. Widdowson, A. Oetting, A. Lakshmanan, H. Cui, N. Hovakimyan, R. F. Wang, A. Kirlik, A. Laviers, and D. Stipanović, "Carebots," *Mechanical Engineering Magazine Select Articles*, vol. 138, no. 09, pp. S8–S13, 2016.
- [5] (2017) Gartner says almost 3 million personal and commercial drones will be shipped in 2017. [Online]. Available: <http://www.gartner.com/newsroom/id/3602317>
- [6] B. A. Duncan and R. R. Murphy, "Comfortable approach distance with small unmanned aerial vehicles," in *IEEE International Conference on Robot and Human Interactive Communication*. IEEE, 2013, pp. 786–792.
- [7] —, "Effects of speed, cyclicity, and dimensionality on distancing, time, and preference in human-aerial vehicle interactions," *ACM Transactions on Interactive Intelligent Systems (TIIS)*, vol. 7, no. 3, p. 13, 2017.
- [8] J. R. Cauchard, K. Y. Zhai, M. Spadafora, and J. A. Landay, "Emotion encoding in human-drone interaction," in *ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2016, pp. 263–270.
- [9] D. Szafir, B. Mutlu, and T. Fong, "Communicating directionality in flying robots," in *ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015, pp. 19–26.
- [10] G. Williams, A. Aldrich, and E. A. Theodorou, "Model predictive path integral control: From theory to parallel computation," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 2, pp. 344–357, 2017.
- [11] C. Widdowson, H.-J. Yoon, V. Cichella, F. Wang, and N. Hovakimyan, "VR environment for the study of collocated interaction between small uavs and humans," in *International Conference on Applied Human Factors and Ergonomics*. Springer, 2017, pp. 348–355.
- [12] T. Marinho, A. Lakshmanan, V. Cichella, C. Widdowson, H. Cui, R. M. Jones, B. Sebastian, and C. Goudeseune, "VR study of human-multicopter interaction in a residential setting," in *IEEE Virtual Reality (VR)*, March 2016, pp. 331–331.
- [13] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of Neuroscience Methods*, vol. 190, no. 1, pp. 80–91, 2010.
- [14] M. C. Mozer, S. Kinoshita, and M. Shettel, "Sequential dependencies in human behavior offer insights into cognitive control," *Integrated Models of Cognitive Systems*, pp. 180–193, 2007.
- [15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B (methodological)*, pp. 1–38, 1977.
- [16] J. B. Rawlings and D. Q. Mayne, *Model predictive control: theory and design*. Nob Hill Pub. Madison, Wisconsin, 2009.
- [17] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, "Aggressive driving with model predictive path integral control," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 1433–1440.
- [18] —, "Information-theoretic model predictive control: Theory and applications to autonomous driving," *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1603–1622, 2018.
- [19] A. Juliani, V.-P. Berges, E. Vckay, Y. Gao, H. Henry, M. Mattar, and D. Lange, "Unity: A general platform for intelligent agents," *arXiv preprint arXiv:1809.02627*, 2018.
- [20] H.-J. Yoon, D. Lee, and N. Hovakimyan, "Hidden Markov model estimation-based q-learning for partially observable Markov decision process," *arXiv preprint arXiv:1809.06401*, 2018.