

# MATHEMATICAL MODELS OF HUMAN DRIVERS USING ARTIFICIAL RISK FIELDS

Emily Jensen<sup>1</sup>, Maya Luster<sup>2</sup>, Hansol Yoon<sup>1</sup>, Brandon Pitts<sup>2</sup>, and Sriram Sankaranarayanan<sup>1</sup>.

**Abstract**—In this paper, we use the concept of artificial risk fields to predict how human operators control a vehicle in response to upcoming road situations. A risk field assigns a non-negative risk measure to the state of the system in order to model how close that state is to violating a safety property, such as hitting an obstacle or exiting the road. Using risk fields, we construct a stochastic model of the operator that maps from states to likely actions. We demonstrate our approach on a driving task wherein human subjects are asked to drive a car inside a realistic driving simulator while avoiding obstacles placed on the road. We show that the most likely risk field given the driving data is obtained by solving a convex optimization problem. Next, we apply the inferred risk fields to generate distinct driving behaviors while comparing predicted trajectories against ground truth measurements. We observe that the risk fields are excellent at predicting future trajectory distributions with high prediction accuracy for up to twenty seconds prediction horizons. At the same time, we observe some challenges such as the inability to account for how drivers choose to accelerate/decelerate based on the road conditions.

## I. INTRODUCTION

We consider the problem of systematically modeling human control actions inside an intelligent transportation system. Ideally, such a model would enable interpretable explanations of why human drivers make certain control decisions in different situations. Moreover, a model of driver decisions should be able to capture the variation in human driving behavior and emulate qualitatively different driving behaviors. Such models of human drivers can be quite helpful in developing autonomous vehicles that behave in a predictable manner and are able to operate on roads with human-driven vehicles [1], [2]. Furthermore, driver models can potentially be used in applications such as runtime monitoring of human drivers to predict dangerous driving behaviors wherein the actions of the driver are “far away” from those expected by our model.

In this paper, we consider probabilistic models of human actions by building upon the concept of *artificial risk fields*. Such risk fields map states of the system to non-negative risk values, wherein larger risk values imply the state is close to a violation. The choice of a control action from a given state by the human operator follows from the risk model in a simple way: the probability that a given control action is chosen is proportional to the exponential of the risk at the state that is reached at a fixed *preview time* by applying that action. We develop this idea in the context of

human control of a car wherein the human operator is tasked with driving the car safely along a road while staying in the designated lane, and at the same time, avoiding obstacles placed on the road. We first show how a family of possible risk functions can be formulated for such tasks, wherein each risk function is obtained by instantiating some unknown parameters to a specific values. We demonstrate how the risk function can yield a probability distribution over possible choices of control input that a human operator may select from a given state, assuming a fixed preview time. We also consider the problem of inferring risk functions from actual human operator data. In particular, we show that deriving maximum likelihood risk function parameters for a class of “additive” risk functions reduces to a convex optimization problem that can be solved to global optimum.

We evaluate the proposed framework on data collected from human drivers inside a simulated driving environment, wherein the humans are tasked to drive the vehicle along a fixed course while avoiding obstacles placed along the vehicle’s path. Using data from six different drivers with up to four trials around the course for each driver, we show that our approach can fit parameters for risk models in each case. We explore the interpretation of these parameters showing how they predict qualitatively different behaviors. Next, we evaluate the ability of our model to predict future trajectories that are close to the ground truth trajectories. Here, we show that our model can provide very accurate predictions with errors that lie within a few meters for predicting the position 20 seconds out into the future. However, at the same time our model is less accurate for predicting how drivers accelerate or decelerate over different portions of their driving tasks.

The main contributions in this paper are as follows: (a) We formalize the risk field-based approach that has been proposed by many researchers in the past [3]–[6]. A key contribution lies in formalizing the driver model based on a risk field as a stochastic model and providing approaches to discovering model parameters from naturalistic data. (b) We instantiate our framework to a driving simulator-based study of human operators driving a vehicle around in a simulated course with obstacles. (c) Our empirical evaluation shows that risk field-based approach can provide reasonable predictions of future trajectories. (d) Finally, we systematically vary risk field parameters to generate distinct driver behaviors.

## II. RELATED WORK

Munir et al. [7] discuss the main challenges facing feedback control with human-in-the-loop; in particular, they discuss the need for developing systematic models of human

\* Support information will go in the final version.

<sup>1</sup> University of Colorado Boulder. Email: first.lastname@colorado.edu

<sup>2</sup> Purdue University. Email: first.lastname@purdue.edu

behavior. Previous approaches to modeling driver behavior rely on cognitive models of human information processing. Salvucci and Gray [8] exploit the tendency of a driver’s gaze to fixate on a near and far point. Subsequent work by Salvucci [9] used models of human declarative and procedural knowledge in the ACT-R cognitive architecture [10] to simulate steering angle and lateral position for navigating curves. Our work also models human operator control choices in a systematic manner. The key differences are two-fold: we our model predicts a distribution over possible control inputs rather than a fixed prediction based on the state. Also, unlike the works mentioned above, we do not aim to model the mental processes that underlie the driver’s decision making.

Other work captures driver behaviors in a qualitative manner. For example, Zhang et al [11] characterized drivers as novices or experts using a pattern recognizer on steering inputs. Similarly, Filev et al [12] used a rule-based system to classify drivers as cautious or aggressive based on the variation in their braking and acceleration behaviors. Finally, Wang et al [13] used k-means clustering to identify key characteristics of long-term driving behaviors such as prudence, stability, conflict proneness, and skillfulness. These approaches aim to develop driver profiles. Our approach can be interpreted similarly by examining the relative weights of the risk model components; we can additionally apply artificial risk fields as a generative model of future behavior.

Recent methods in modeling operator behavior are based on navigating “interaction fields” in the task environment [5]. Foundational work by Gibson et al [4] hypothesizes that humans navigate a “field of safe travel” by evaluating possible paths based on subjective experience and objective physical limitations. In the recent work of Kolekar et al [14], participants in a driving simulator were asked to react to obstacles placed at varying positions relative to their vehicle. Based on recorded reactions, the authors constructed a “driver’s risk field” surrounding the vehicle. In a subsequent work [6], they then quantified a driver’s perceived risk as the product of their risk field and the cost of certain events (colliding with obstacles). This leads to a controller which generates human-like behavior in a variety of scenarios when set to maintain risk under a certain threshold. The motivation of this paper is similar to the work of Kolekar et al [6], [14] in that we seek an interpretable and generative model of driving behavior grounded in theories of human reasoning and decision making. Our approach differs from the above work in several important respects. First, we define a risk field as a characteristic of the task environment and control inputs selected by the operator. The operator then stochastically navigates this risk space with the goal of minimizing risk. A second distinction is that because the risk fields presented here are defined in the task environment, they extend to other scenarios besides driving.

Our approach is closely related to *inverse reinforcement learning* where the vehicle model and operator’s actions are captured by a Markov Decision Process (MDP) model with unknown reward functions. The goal is to infer these

unknown rewards either through solving an optimization problem [15], [16] or through Bayesian methods [17]. There has been a long history of using inverse reinforcement to explain the actions of human operators inside a known environment [18]. The recent work of Ozkan et al studies how inverse reinforcement learning can be used to learn a driver model that is able to predict lead vehicle following behaviors of human drivers in a 3D driving simulation environment [19]. Our approach bears many similarities to IRL: for instance, we can view risk fields as a (negative) “reward” function that the driver is minimizing. However, some key differences exist: we explicitly consider a “preview time” that the operator looks ahead into the future. This allows us to keep our risk functions simple since they apply to the state that is reached at some time in the future. Inverse reinforcement learning approaches compute rewards/risks that apply to the current state. This means that they have to consider more complicated functions than we do. As a result of our setup, we also have the benefit of solving a convex optimization problem and thus guarantee that we can compute the most likely model.

Rather than modeling the driver’s risk perception and control choice, data-driven approaches such as Kim et al [20] and Long et al [21] train recurrent neural networks that input numerous features such as the vehicle’s past trajectories and from its surrounding environment to predict the future trajectories of the vehicle. While these approaches are promising as predictors of future trajectories, they require a larger volume of data to reliably train and test a recurrent neural network. It is often challenging to interpret variations between drivers or in general understand these models once they are trained. Nevertheless, data driven approaches have proven more versatile and capable of handling many more situations than our approach in this paper. Our work is currently aimed towards more narrowly defined settings although we hope to generalize it in future iterations of our framework to handle a richer variety of driving scenarios.

### III. ARTIFICIAL RISK FIELDS

We describe a general approach to defining human operator behavior using artificial risk fields. Subsequently, we will apply the framework to a driving task in Section IV.

#### A. Problem Formulation

We consider a single vehicle inside a known environment of *states*  $\mathbf{x} \in X$ . The human driver’s task is to control the vehicle from a starting configuration  $A$  to a goal configuration  $B$ . Additionally, we designate a set of *obstacles*  $O_1, \dots, O_m$ ; each obstacle  $O_i \subseteq X$  represents an unsafe configuration. It may also be natural to specify a “desired” path  $\pi$  that connects the start to the end state, that the vehicle should stay close to (Figure 1).

The vehicle is modeled by its *dynamics*:  $\frac{d\mathbf{x}}{dt} = f(\mathbf{x}(t), \mathbf{u}(t))$ , wherein  $\mathbf{x}(t)$  models the state at some time  $t$  and  $\mathbf{u}(t) \in U$  models the action (control input) at time  $t$  and  $U$  is the set of actions available to the human driver. The function  $f$  is assumed to be a fixed and known state update

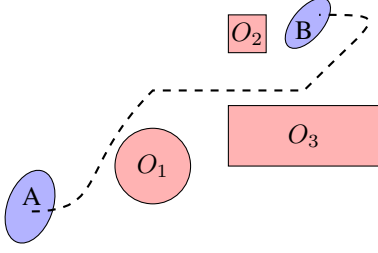


Fig. 1. Schematic diagram of the overall task setup showing initial set of states  $A$ , target set  $B$ , obstacles  $O_i$  and optionally a reference path.

function. Our approach makes some key assumptions about the behavior of the human operator:

- 1) The operator knows the state  $\mathbf{x}$ , or at least, those state variables involved in choosing the control.
- 2) The operator model is Markovian — i.e, the probability distribution depends on the current state  $\mathbf{x}$  and not necessarily on the path taken to reach the state.

Whereas the assumptions above are somewhat restrictive, we note that our goal is to build a model that predicts the operator’s decision making rather than capture the mental processes involved in the decision making.

The overall goal of this paper is to predict what control actions are likely to be chosen by the human operator at a given state. That is, we seek to model the probability distribution of  $\mathbb{P}(\mathbf{u}|\mathbf{x})$  that an action  $\mathbf{u}$  that is chosen by the human operator at state  $\mathbf{x}$ . Our model makes the following assumptions about the operator’s control selection strategy:

- 1) Each state  $\mathbf{x} \in X$  is associated with a non-negative *risk value*  $\text{risk}(\mathbf{x}; \mathbf{p})$  which provides an aggregate numerical score, wherein  $\mathbf{p}$  denotes a set of parameters that may be specific to an individual operator at a given time. The higher the risk score associated with a state, the closer it is to being a property violation such as entering a forbidden obstacle region or deviating too far from a desired path.
- 2) The operator plans ahead to some “preview” time  $\delta_p > 0$  into the future.
- 3) The operator’s decision making balances two factors: the risk of the future state that would be reached if a particular control were chosen against the magnitude of the control input. Thus, the operator would prefer not to apply extreme values of brakes/acceleration or steering inputs while at the same time they would prefer to stay away from obstacles and close to the center of their designated lane.

We will first describe each component of the model starting with the risk function. Next, we will describe how the overall probability distribution is defined.

## B. Risk Function

The risk may be defined by many factors including the proximity of the state to various obstacles and the deviation of the state from the desired path  $\pi$ . The risk function

$\text{risk}(\mathbf{x}; \mathbf{p})$  is given by:

$$\sum_{j=1}^m p_j \text{obstacleRisk}(\mathbf{x}, O_j) + p_{m+1} \text{deviationRisk}(\mathbf{x}, \pi),$$

wherein  $\text{obstacleRisk}(\mathbf{x}, O_j)$  is a function that measures the risk connected with the state  $\mathbf{x}$  being inside (or close to) the obstacle  $O_j$ , and  $\text{deviationRisk}(\mathbf{x}, \pi)$  measures the risk arising from the state  $\mathbf{x}$  being far away the reference path  $\pi$  (if one is given in the problem formulation). In general, any function that ensures that the risk is monotonically decreasing as one moves away from the obstacle can be chosen. Similarly,  $\text{deviationRisk}(\mathbf{x}, \pi)$  will be 0 if  $\mathbf{x}$  lies on the reference path, and increases monotonically as the distance from the state  $\mathbf{x}$  to the reference path  $\pi$  increases.

Finally, we note that the parameters  $\mathbf{p} : (p_1, \dots, p_{m+1})$  are non-negative weights that model the relative weightages associated with avoiding various obstacles and being close to the reference path. The choice of these parameters will affect the nature of the risk function. In Section V, we demonstrate how parameters for a risk model are chosen given observed experimental data.

## C. Overall Operator Model

The next component of the risk model concerns the assumption of a preview time. Let  $\mathbf{x}$  be a current state and  $\mathbf{u}$  be a control action under consideration. We assume that the operator computes the state  $\mathbf{x}'$  at some fixed time  $\delta_p > 0$  in the future. In other words, let  $\mathbf{x}'(\mathbf{u}, \delta_p)$  be the state that results at time  $t + \delta_p$  if the control action  $\mathbf{u}$  were chosen at time  $t$  and held constant. Also, we associate a non-negative cost to each control action  $\mathbf{u}$  denoted by  $\text{cost}(\mathbf{u}; \mathbf{q})$ . Once again, the cost model can be parameterized by a set of unknown parameters  $\mathbf{q}$  that will be estimated from experimental data.

The operator model we formulate assumes that

$$\mathbb{P}(\mathbf{u}|\mathbf{x}) \propto \exp(-\text{risk}(\mathbf{x}'(\mathbf{u}, \delta_p); \mathbf{p}) - \text{cost}(\mathbf{u}; \mathbf{q})).$$

Suppose the set of possible actions  $U$  is a finite set  $\{\mathbf{u}_1, \dots, \mathbf{u}_N\}$ , then we write the exact expression as Eq. (1). The denominator normalizes the probability over all actions. For continuous set of control actions, we can replace the summation by an integral over the set  $U$ . Doing so, we obtain the following expression for  $\mathbb{P}(\mathbf{u}|\mathbf{x})$ :

$$\frac{\exp(-\text{risk}(\mathbf{x}'(\mathbf{u}, \delta_p); \mathbf{p}) - \text{cost}(\mathbf{u}; \mathbf{q}))}{\sum_{j=1}^N \exp(-\text{risk}(\mathbf{x}'(\mathbf{u}_j, \delta_p); \mathbf{p}) - \text{cost}(\mathbf{u}_j; \mathbf{q}))}. \quad (1)$$

The operator model implicitly assumes that (a) the operator can forecast a future state  $\mathbf{x}'(\mathbf{u})$  some time  $\delta_p$  in the future as a result of a control input  $\mathbf{u}$ ; and (b) chooses control actions which yield future states with lower risk + cost values preferentially over those with higher risk + cost values.

## IV. DRIVING TASK

We describe the driving task that will be the central case-study to motivate our work and develop a risk field model specific to the task.

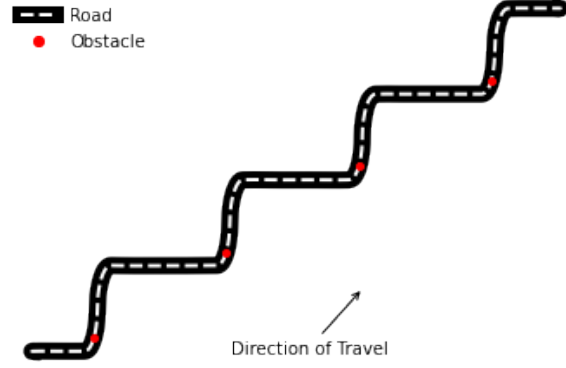


Fig. 2. (Left) Picture of the NADS miniSim setup showing a participant driving along a course (daytime simulation), (Right) plot of the centerline of the simulated course showing obstacle placement as red circles.

### A. Task Description

The driving task is performed in a medium-fidelity driving simulation environment developed by the National Advanced Driving Simulator (NADS miniSim) at Purdue University [22]. The system includes three high resolution monitors for displaying the driving environment and a smaller monitor for the vehicle dashboard display. The user controls a steering wheel and foot pedals for acceleration and braking as in a standard automobile (Figure 2, left).

*Driving Scenario.* The driving scenario consists of driving the simulated vehicle at night time on a two lane city highway with four obstacles placed along the route. Illumination using street lights were present. The overall simulated driving course distance was roughly 4.8 km (3 miles). To increase the difficulty of the task, participants were asked to drive one handed with their non dominant hand. There were no oncoming, leading, or trailing vehicles. The obstacles were placed so that they were visible only after the participant rounded the curve (Figure 2, right).

The objectives for the human driver are as follows:

- 1) The operator must practice safe driving by keeping within their lane and minimizing deviations. They must never exit the paved road.
- 2) Obstacles (a tire) placed in the operator's lane are to be avoided.
- 3) Vehicle speed is to be maintained as close as possible to 45 mph ( $\approx 20$  m/s) at all times.

*Participants.* The study was conducted with six participants (3 male, 3 female) with a mean age of 21.33 years (SD = 0.82). Participants were all undergraduate students at Purdue University<sup>1</sup>, and were all engineering senior undergraduate students. On average, the participants had 4.2 years of driving experience, with all of them reporting having driven 10K or more miles per year, on the average. The participants were allowed to practice driving the vehicle on the simulator using a daytime practice course that involved an open highway.

*Data Collection.* Each participant drove the course over three (or in one case, four) separate trials, yielding nineteen separate trials for the six participants, in total. Data collected includes the position, velocity, heading angle, steering wheel position, accelerator/brake pedal positions sampled at 60 Hz.

### B. Risk Field Formulation

We will now derive risk models for the human driving task. First, we will describe a simple unicycle model for the vehicle's dynamics. This model is appropriate since effects such as cornering over tight turns, wheel slip and skids are not important for the speed and road conditions that were simulated in the study. The state of the vehicle is described by  $\mathbf{x} : (x, y, v, \psi)$ , wherein  $x, y$  denote the position in a fixed coordinate frame,  $v$  describes the velocity of the vehicle and  $\psi$  is the heading angle. The control inputs are  $u_1$ : the acceleration (or deceleration) and  $u_2$ : the turning rate. The dynamics are described by the ODEs:

$$\left. \begin{aligned} \dot{x} &= v \cos(\psi) \\ \dot{y} &= v \sin(\psi) \\ \dot{v} &= u_1 \\ \dot{\psi} &= u_2 \end{aligned} \right\} \quad (2)$$

We define the function  $\text{ptLineDistance}((x, y), C)$  as the Euclidean distance from a given position  $(x, y)$  to the nearest point in the center-line  $C$ .

Similarly, we are given a list of obstacle positions  $O : [(x_{o,1}, y_{o,1}), \dots, (x_{o,4}, y_{o,4})]$ . Each obstacle has a fixed diameter  $d_o = 0.3$  meters. We define the function  $\text{obstacleDistance}((x, y), O)$  as the Euclidean distance from a given position  $(x, y)$  to the obstacle that will be encountered next in the vehicle's direction of travel.

The overall risk for a given state  $\mathbf{x} : (x, y, v, \psi)$  and control  $\mathbf{u}$  is given by  $\text{risk}(\mathbf{x})$ :

$$\text{risk}(\mathbf{x}) : \left\{ \begin{aligned} &\mathbf{A} \cdot \text{ptLineDistance}((x, y), C)^2 + \\ &\mathbf{B} \cdot \exp\left(-\frac{\text{obstacleDistance}((x, y), O)^2}{d_o^2}\right) + \\ &\mathbf{C} \cdot (v - v_{tgt})^2 \end{aligned} \right. \quad (3)$$

and the cost of the control input is given by  $\text{cost}(\mathbf{u})$ :

$$\text{cost}(\mathbf{u}) : \mathbf{D} \cdot u_1^2 + \mathbf{E} \cdot u_2^2. \quad (4)$$

<sup>1</sup>This study was approved by Purdue IRB number 1905022220

**Data:** risk, cost: risk/cost functions,  $\mathbf{x}_0$ : Initial State,  $\delta_p$ : preview time,  $\delta$ : time step,  $n_s$ : number of simulation steps,  $U : \{\mathbf{u}_1, \dots, \mathbf{u}_N\}$  all control inputs  
**Result:** Sample Trajectory:  $\mathbf{x}(0), \dots, \mathbf{x}(n_s\delta)$   
 $\mathbf{x}(0) \leftarrow \mathbf{x}_0$ ;  
**for**  $s \leftarrow 1, \dots, n_s$  **do**  
  **for** *each*  $\mathbf{u}_j \in U$  **do**  
    /\* Simulate until the preview time. \*/  
     $\mathbf{x}'_j \leftarrow \text{nextState}(\mathbf{x}(\delta(s-1)), \mathbf{u}_j, \delta_p)$ ;  
    /\* Calculate Risk. \*/  
     $p(\mathbf{u}_j) \leftarrow \exp(-\text{risk}(\mathbf{x}'_j; A, B, C) - \text{cost}(\mathbf{u}_j; D, E))$ ;  
  **end**  
  sample  $\mathbf{u} \in U$  with probability  $p(\mathbf{u}) / \sum_{k=1}^N p(\mathbf{u}_k)$ ;  
   $\mathbf{x}(\delta s) \leftarrow \text{nextState}(\mathbf{x}, \mathbf{u}, \delta)$ ; /\* State for  $\delta s$  \*/  
**end**

**Algorithm 1:** Algorithm for sampling a trajectory given risk and cost functions, initial states.

Here  $A, \dots, E \geq 0$  are unknown parameters whose values will determine the actual tradeoffs that the driver makes while staying in their lane and avoiding the obstacles during the execution of the task.

We consider control inputs  $u_1 \in \{-1, -0.9, \dots, 0.9, 1\}$  (units are  $\text{m/s}^2$ ) and  $u_2 \in \{-0.5, -0.45, \dots, 0.45, 0.5\}$  (units are  $\text{radians/s}$ ), yielding 400 discrete choices for  $(u_1, u_2)$ . For a given state  $\mathbf{x}$ , the probability of control inputs  $(u_1, u_2)$  being chosen  $\mathbb{P}((u_1, u_2) | \mathbf{x})$  is described once again by Eq. (1). Here we define the risk and costs by Eqs. (3) and (4). The next state  $\mathbf{x}'(\mathbf{u}, \delta_p)$  is obtained by simulating the ODE in Eq. (2).

Algorithm 1 shows the overall algorithm for sampling a trajectory from the risk model.

## V. MODEL FITTING

### A. Maximum Likelihood Estimation

In this section, we consider how to infer a risk field given data in the form of states  $\mathbf{x}(t)$  and controls  $\mathbf{u}(t)$ . We will assume that the risks and costs are *additive* over component functions as follows:

$$\text{risk}(\mathbf{x}; \mathbf{p}) : \sum_{j=1}^m p_j f_j(\mathbf{x}), \quad \text{cost}(\mathbf{u}; \mathbf{q}) : \sum_{i=1}^l q_i g_i(\mathbf{u}). \quad (5)$$

Note however, that we do not assume much for functions  $f_j, g_i$  other than that they are non-negative and well-defined over the relevant values of  $\mathbf{x}, \mathbf{u}$ . The parameters for risk and cost functions are collected as a vector  $(p_1, \dots, p_m, q_1, \dots, q_l)$ . Assuming that the controls are chosen from a finite set  $U : \{\mathbf{u}_1, \dots, \mathbf{u}_N\}$ , fixing  $\delta_p$  to be the preview time and  $\text{nextState}(\mathbf{x}, \mathbf{u}, \delta_p)$  being the state reached starting from current state  $\mathbf{x}$  if control  $\mathbf{u}$  is applied for time  $\delta_p$ . Recall that the model chooses a control input  $\mathbf{u}$  for a state  $\mathbf{x}$  in proportion to the risk and cost according to Eq. (1).

Let us assume that we are given driving data of the form  $(\mathbf{x}(t_i), \mathbf{u}(t_i))$  consisting of states and controls applied at various times  $t_i$  for  $i = 1, \dots, M$ . Our goal is to find

risk parameters  $\mathbf{p}, \mathbf{q}$  for Eq. (5) that maximizes the overall log-likelihood  $\mathcal{L}(\mathbf{p}, \mathbf{q}) : \sum_{i=1}^M \log \mathbb{P}(\mathbf{u}(t_i) | \mathbf{x}(t_i))$ , wherein  $\mathbb{P}(\mathbf{u}(t_i) | \mathbf{x}(t_i))$  is as given in Eq. (1).

Note that if the risk and cost models are additive as in Eq. (5), then the overall log-likelihood  $\mathcal{L}(\mathbf{p}, \mathbf{q})$  is a *concave function* for a fixed value of  $\delta_p$ . This means that we can solve the maximization problem of a concave function (or alternatively minimization of a convex function) to obtain a global optimum using standard off-the-shelf convex optimization tools [23].

**Theorem 1:** If the risk and cost models are additive as in Eq. (5), then the overall log-likelihood  $\mathcal{L}(\mathbf{p}, \mathbf{q})$  is a *concave function* for a fixed value of  $\delta_p$ .

**Proof:** The proof consists in observing that  $\log \mathbb{P}(\mathbf{u} | \mathbf{x})$  is concave function of  $\mathbf{p}, \mathbf{q}$ . Let  $\mathbf{x}'(\mathbf{u})$  denote the value of  $\text{nextState}(\mathbf{x}, \mathbf{u}, \delta_p)$ . Expanding Eq. (1) using the form of the risk model in (5), we obtain an expression for  $\log \mathbb{P}(\mathbf{u} | \mathbf{x})$  in Eq. (6)

Since  $\mathbf{x}, \mathbf{u}$  are given to us in the data, the terms  $f_j(\mathbf{x}'(\mathbf{u}))$  and  $g_i(\mathbf{u})$  are all fixed constants. Thus, as a function of  $\mathbf{p}, \mathbf{q}$ , we note that  $\log \mathbb{P}(\mathbf{u} | \mathbf{x})$  is the difference of a linear function over  $\mathbf{p}, \mathbf{q}$  and the log-sum-exp of linear function over  $\mathbf{p}, \mathbf{q}$ . This is a difference of a concave function and a convex function, which is itself concave. The overall likelihood is the sum of concave functions, and is concave. ■

### B. Fitting Parameters From Obstacle Avoidance Data

In this section, we report on the application of the maximum likelihood minimization approach to the data obtained from six human drivers in the NADS vehicle simulator, as described in Section IV.

We recall that each participant drove along a road with obstacles placed at periodic intervals. In particular, each “trial” by a participant involved four encounters with the obstacle. We will fit the risk model parameters using the data from each obstacle, using the `scipy.optimize` module for various values of  $\delta_p \in \{0.6, 0.8, 1.0, 1.2\}$  seconds. The risk functions used are described in Section IV-B and in particular Eqs. (3) and (4). This yielded  $19 \text{ trials} \times 4 \text{ obstacles} = 76$  fit models for each of the four  $\delta_p$  values.

For each obstacle encounter, we selected the preview time  $\delta_p$  which maximizes the likelihood of the data. Of the four  $\delta_p$  values considered, 94.7% of fitted models achieved a maximum likelihood using  $\delta_p = 1.2$  seconds. Thus, for a car driven at  $20 \text{ m/s}$ , the preview distance is 24 meters.

Table I shows the distribution of fit parameters as the median value as well as extremely low (5th percentile) and extremely high (95th percentile) values. We see that each parameter takes on a different range of values, with the parameter  $C$  (associated with staying close to the target velocity of  $20 \text{ m/s}$ ) varying very little ( $0 - 0.025$ ), whereas  $B$  (associated with the weightage placed on obstacle avoidance) encompasses a wide range ( $0 - 110.86$ ).

$$\log \mathbb{P}(\mathbf{u} | \mathbf{x}) : \left\{ -\sum_{j=1}^m p_j f_j(\mathbf{x}'(\mathbf{u})) - \sum_{i=1}^l q_i g_i(\mathbf{u}) - \log \left( \sum_{k=1}^N \exp \left( -\sum_{j=1}^m p_j f_j(\mathbf{x}'(\mathbf{u}_k)) - \sum_{i=1}^l q_i g_i(\mathbf{u}_k) \right) \right) \right\} \quad (6)$$

TABLE I

DISTRIBUTION OF PARAMETERS FIT AROUND ONE OBSTACLE, USING THE BEST PREVIEW TIME  $\delta_p$ . THERE WERE 76 TOTAL FIT MODELS.

Quantile	A	B	C	D	E
5%	0.248	0.000	0.000	0.000	14.233
50%	0.544	16.349	0.000	1.416	40.782
95%	0.939	110.864	0.025	11.827	99.543

TABLE II

DEVIATION (METERS) OF GENERATED TRAJECTORY FROM ACTUAL HUMAN TRAJECTORY, ACROSS ALL SUCCESSFUL COURSE TRIALS. RESULTS ARE REPORTED AT DIFFERENT TIMES FROM THE STARTING POSITION (SECONDS)

	1s	2s	5s	10s	20s
min	0.002	0.051	0.002	0.180	0.104
median	0.064	0.234	0.493	1.114	0.764
max	0.245	0.944	1.655	2.373	1.482

## VI. EVALUATING DRIVER MODELS

We first provide a preliminary analysis to evaluate the accuracy of our method for predicting driver trajectories. Of the 19 initial recorded course trials, we removed any where the driver collided with an obstacle, yielding 17 successful trials. For each successful trial, we fit risk field parameters using the formulation in Section IV-B and the driver data from the first two obstacles in the trial. Using these parameters, we used Algorithm 1 to generate 100 trajectories for the held out data of the last two obstacles in the trial. We used a preview time  $\delta_p = 1.2$  for all of the trajectories based on the analysis from Section V-B.

To define a single trajectory for comparison with the actual driver behavior, we took the median  $x$  and  $y$  value over the 100 trajectories at each time point. We then defined the divergence of the generated trajectory as the distance from the median point to the line created by the human trajectory. Table II shows the minimum, median, and maximum divergence from the human trajectory at times  $t \in \{1, 2, 5, 10, 20\}$  seconds from the initial position.

Table II shows that, like one would intuitively expect, the deviation from the actual human trajectory increases over time, except between 10 and 20 seconds. The change in deviations may be a function of the course characteristics (e.g., rounding a turn) and also show that our model can self-correct based on the high-level priorities defined in the risk field model. Additionally, these results show that the model is able to predict the future position 20 seconds ahead with an error of less than 3 meters. This is promising, given that the lane width in the driving task was 3 meters.

Figure 3 shows sample  $(x, y)$  trajectories predicted by our model and velocities over time for three separate initial conditions drawn from the actual driver data. We also plot the actual “ground-truth” data for each of these situations. It is interesting to see that the simulated  $(x, y)$  trajectories are viable trajectories that keep close to the center line while

avoiding obstacles. In the bottom row of Figure 3, we see that the predicted velocity deviates from the true human driver velocity by as much as 4 m/s. The mean absolute difference in predicted velocity versus actual velocities are 0.1 m/s for predicting 1 seconds out into the future, 1.3 m/s for 10 second prediction horizons and 2.5 m/s for 20 second horizon. However, we also observe that our model has the tendency to under-estimate the actual velocity around sharp turns: it is likely that the driver allows the vehicle to move towards the edge of their lane to reduce steering effort and allow themselves to accelerate. We conclude that the participants do not prioritize the instruction to maintain their velocity around 20 m/s, while focusing more on maintaining their lane position and avoiding obstacles. Modeling their choice of velocities requires considerations that are subtly different from their perceived risk such as their self-confidence.

## VII. CHARACTERIZING DRIVER BEHAVIOR

Showing the overall accuracy of our model, we reach the main research question, do the risk model parameters account for different types of obstacle avoidance behavior? To answer this, we will visualize generated trajectories using different parameter configurations. For each condition, we used Algorithm 1 to generate 20 trajectories around the course segment for the first obstacle. Our baseline comparison uses the median value for each parameter when calculating the risk field. To simulate the condition using the low and high values of a parameter, we used the 5th and 95th percentiles of the parameter, respectively, leaving the remainder of the parameters at their median level (see Table I). We used a preview time of  $\delta_p = 1.2$  as in the previous section.

Figure 4 shows the differences as we vary parameters from the baseline. In 4a we see that high values for  $A$  lead to trajectories very close to the centerline while low values for  $A$  stray farther from the centerline. While 4b shows more consistent trajectories between conditions, higher values for  $B$  leave a higher margin when passing the obstacle compared to the lower value condition.

Since parameter  $D$  impacts user controls, we visualize the acceleration under the different conditions rather than the physical position. In Figure 4c we see that high values of  $D$  lead to much more consistent acceleration compared to the low values. For parameter  $E$ , in 4d low value trajectories show a sharper decrease in steering rate after 5 time steps compared to the baseline and high level trajectories.

Overall, we conclude that varying the risk model parameters has the expected change in the trajectories. For instance, increasing  $B$  causes the trajectories to clear the obstacle with a much larger safety margin. Increasing  $A$  on the other hand has the opposite effect of bringing trajectories closer to the centerline.

## VIII. DISCUSSION AND FUTURE WORK

In this paper, we have presented an approach to model control choices of the human driver by quantifying the risk and showing how the risk model can be inferred from data. We have also demonstrated our approach on actual human



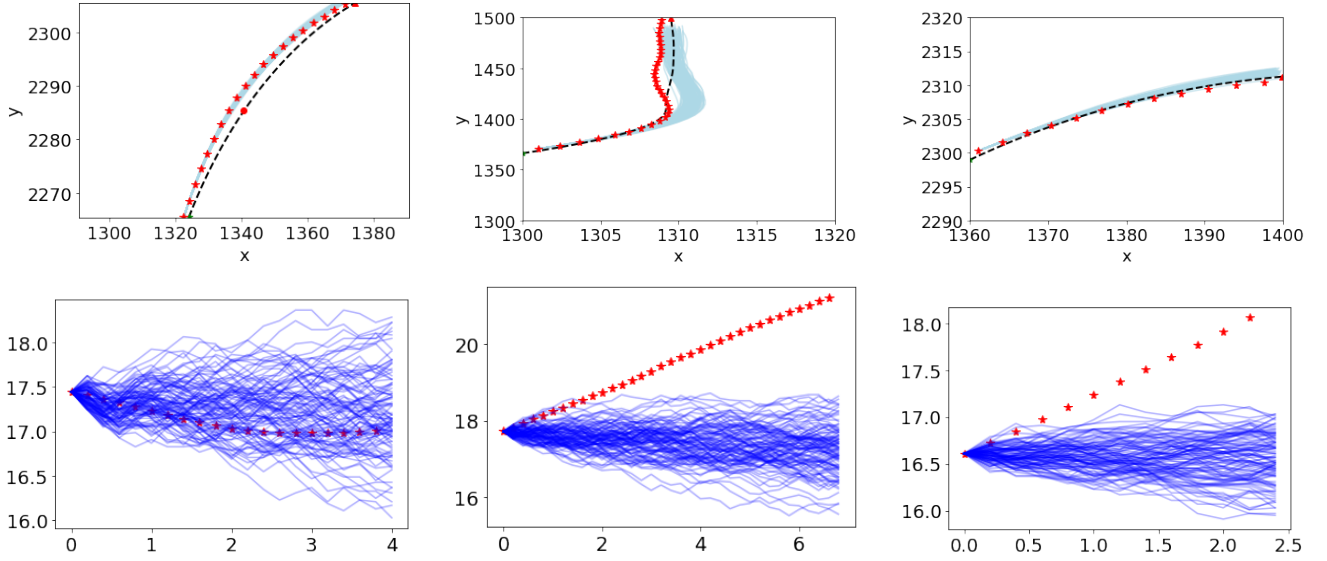


Fig. 3. (Top Row) Sample  $(x, y)$  trajectories generated by the risk model against ground truth shown by red stars with centerline shown as a dashed black line and obstacle shown as red circle. **Warning:**  $x$  and  $y$  axes are drawn to different scales. (Bottom Row) Corresponding velocity (m/s) values over time against ground truth.

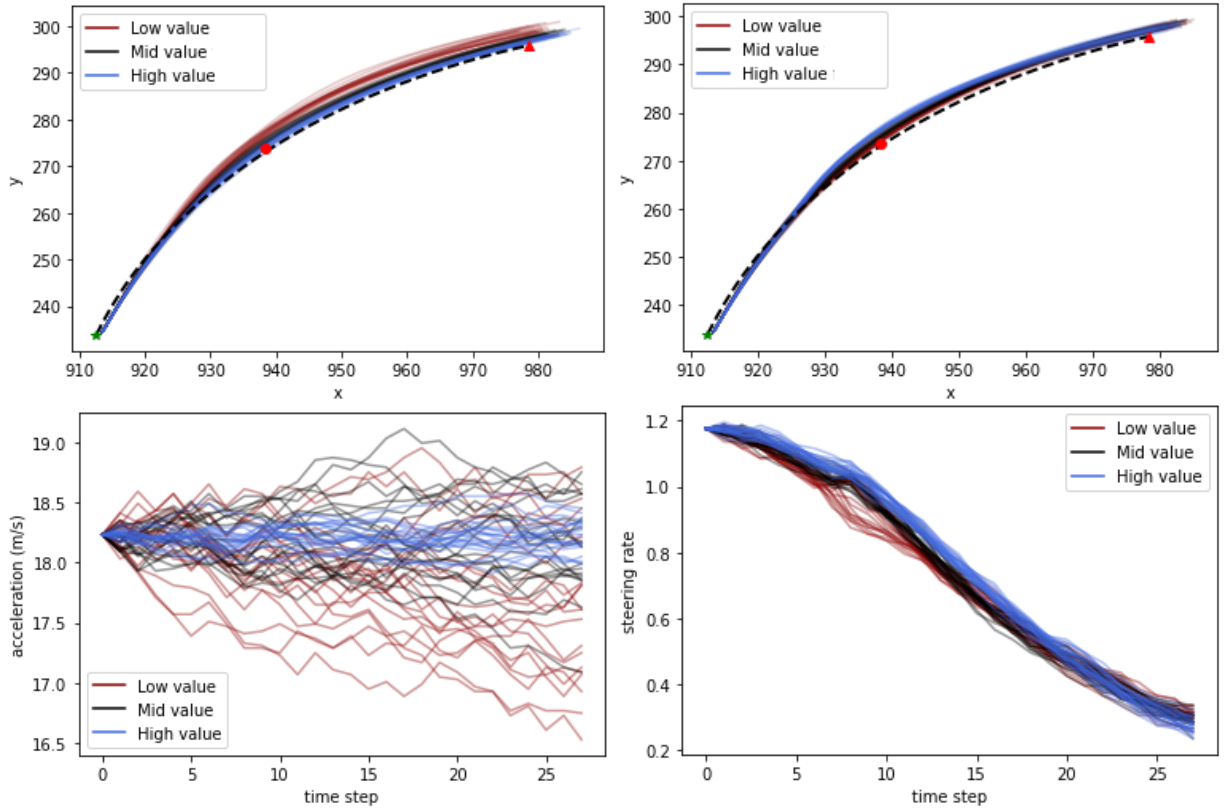


Fig. 4. Simulated trajectories using extreme low (5<sup>th</sup> percentile) and high values (95<sup>th</sup> percentile) for the parameters. From left to right, top to bottom; (a) parameter  $A$ : center line deviation risk weightage; (b) parameter  $B$ : obstacle avoidance risk weightage; (c) parameter  $D$ : cost for acceleration; (d) parameter  $E$ : cost for turning rate control.

driving data from a medium-fidelity simulation environment showing that our models can accurately predict future positions and generate qualitatively different driving behaviors. In particular, we show that deviation of generated trajectories

from the human trajectory remains relatively stable over time periods up to 20 seconds into the future.

The main area for improvement is that our model currently does not capture how human operators control the velocity.

We plan to improve this aspect of our model in our future work. For example, we can consider more complex representations of risk and cost beyond the simple quadratic model presented here. Additionally, a driver's choice of velocity may depend on other factors such as their confidence in driving or the overall level of risk of the current situation. The fact that our models had very small values for the  $C$  parameter that measures velocity deviations from intended target indicates that human driver behavior during the task may have been influenced by factors different from risk. While our model was defined to maintain a predefined velocity as stated in the task instructions, we observed that the drivers themselves did not adhere to this requirement.

The main result of this paper shows that by using this risk model framework with simple models for risk and control cost, we are able to generate distinct driver behaviors such as obstacle avoidance and keeping to the center of the lane. Using real driver data collected in a simulation environment, we have also shown that we can extract unique parameters that characterize individual driver behavior. Future work should investigate how accurately these models track more complex human behavior over time. Additionally, this model can be used as part of a predictive run-time monitoring system, where the goal is to predict impending violations of safety property (i.e., colliding with an obstacle) ahead of time. This system could be integrated into future driver safety interfaces and be used to study potential handover protocols with autonomous driving subsystems.

At a general level, this framework can be used to model a variety of scenarios and adapted to test hypotheses about human operator behavior. As discussed above, we noted that the participants in the study did not maintain the target velocity given in the task instructions. When fitting the risk model parameters, this behavior was indicated by the fact that the fitted values for parameter  $C$  were heavily skewed to 0, indicating no effect. Future work can systematically test different forms of the risk function to see which is a better fit for human behavior. This framework may also be adapted in an attempt to infer the human's true reward function during the driving task.

In the future, we also plan to build on this framework in order to address more dynamic scenarios involving multiple vehicles and moving obstacles. Finally, we plan to address more complex task requirements in our framework.

## REFERENCES

- [1] C. Basu, Q. Yang, D. Hungerman, M. Singhal, and A. D. Dragan, "Do You Want Your Autonomous Car To Drive Like You?" in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. New York, NY, USA: ACM, mar 2017, pp. 417–425. [Online]. Available: <https://dl.acm.org/doi/10.1145/2909824.3020250>
- [2] L. Sun, W. Zhan, M. Tomizuka, and A. D. Dragan, "Courteous Autonomous Cars," in *International Conference on Intelligent Robots (IROS) 2018*, aug 2018. [Online]. Available: <http://arxiv.org/abs/1808.02633>
- [3] Y. Rasekhipour, "Prioritized Obstacle Avoidance in Motion Planning of Autonomous Vehicles," Ph.D. dissertation, University of Waterloo, 2017.
- [4] J. J. Gibson and L. E. Crooks, "A theoretical field-analysis of automobile-driving," *The American Journal of Psychology*, vol. 51, no. 3, pp. 453–471, 1938. [Online]. Available: <http://www.jstor.org/stable/1416145>
- [5] E. E. Kadar and R. E. Shaw, "Toward an ecological field theory of perceptual control of locomotion," *Ecological psychology*, vol. 12, no. 2, pp. 141–180, 2000.
- [6] S. Kolekar, J. de Winter, and D. Abbink, "Human-like driving behaviour emerges from a risk-based driver model," *Nature Communications*, vol. 11, no. 1, 2020. [Online]. Available: <http://dx.doi.org/10.1038/s41467-020-18353-4>
- [7] S. Munir, J. A. Stankovic, C. J. M. Liang, and S. Lin, "Cyber physical system challenges for human-in-the-loop control," in *8th International Workshop on Feedback Computing*, 2013.
- [8] D. D. Salvucci and R. Gray, "A two-point visual control model of steering," *Perception*, vol. 33, no. 10, pp. 1233–1248, 2004.
- [9] D. D. Salvucci, "Modeling driver behavior in a cognitive architecture," *Human Factors*, vol. 48, no. 2, pp. 362–380, 2006. [Online]. Available: <https://doi.org/10.1518/00187200677724417>
- [10] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An Integrated Theory of the Mind," *Psychological Review*, vol. 111, no. 4, pp. 1036–1060, 2004. [Online]. Available: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-295X.111.4.1036>
- [11] Y. Zhang, W. C. Lin, and Y.-K. S. Chin, "A pattern-recognition approach for driving skill characterization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 905–916, 2010.
- [12] D. Filev, J. Lu, K. Prakah-Asante, and F. Tseng, "Real-time driving behavior identification based on driver-in-the-loop vehicle dynamics and control," in *IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 2009, pp. 2020–2025.
- [13] J. Wang, M. Lu, and K. Li, "Characterization of longitudinal driving behavior by measurable parameters," *Transportation Research Record*, no. 2185, pp. 15–23, 2010.
- [14] S. Kolekar, J. D. Winter, and D. Abbink, "Which parts of the road guide obstacle avoidance? Quantifying the driver's risk field," *Applied Ergonomics*, vol. 89, no. July, 2020. [Online]. Available: <https://doi.org/10.1016/j.apergo.2020.103196>
- [15] A. Y. Ng and S. Russell, "Algorithms for inverse reinforcement learning," in *Proc. International Conf. on Machine Learning (ICML)*. Morgan Kaufmann, 2000, pp. 663–670.
- [16] B. D. Ziebart, A. L. Maas, J. A. Bagnell, A. K. Dey, et al., "Maximum entropy inverse reinforcement learning," in *Aaai*, vol. 8. Chicago, IL, USA, 2008, pp. 1433–1438.
- [17] D. Ramachandran and E. Amir, "Bayesian inverse reinforcement learning," in *AAAI*, 01 2007, pp. 2586–2591.
- [18] C. L. Baker, R. Saxe, and J. B. Tenenbaum, "Action understanding as inverse planning," *Cognition*, vol. 113, pp. 329–349, 2009.
- [19] M. F. Ozkan, A. J. Rocque, and Y. Ma, "Inverse reinforcement learning based stochastic driver behavior learning," *CoRR*, 2021, cf. <https://arxiv.org/abs/2107.06344>.
- [20] B. Kim, C. M. Kang, J. Kim, S. H. Lee, C. C. Chung, and J. W. Choi, "Probabilistic vehicle trajectory prediction over occupancy grid map via recurrent neural network," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 399–404.
- [21] L. Xin, P. Wang, C.-Y. Chan, J. Chen, S. E. Li, and B. Cheng, "Intention-aware long horizon trajectory prediction of surrounding vehicles using dual lstm networks," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 1441–1446.
- [22] "National advanced driving simulator (nads) minisim," <https://nads.uiowa.edu/minisim>, accessed: 2022-Mar-01.
- [23] S. Boyd and S. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.