Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Learning-Based Risk-Averse Model Predictive Control for Adaptive Cruise Control with Stochastic Driver Models ⋆

Mathijs Schuurmans * **Alexander Katriniok** **
Hongtei Eric Tseng *** **Panagiotis Patrinos** *

* *Department of Electrical Engineering* ESAT-STADIUS, *KU Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium (e-mail: {mathijs.schuurmans, panos.patrinos}@esat.kuleuven.be).*
** *Ford Research & Innovation Center, 52072 Aachen, Germany (e-mail: de.alexander.katriniok@ieee.org)*
*** *Research & Innovation Center, Ford Research Laboratories, Dearborn, MI 48124 USA*

Abstract
We propose a learning-based, distributionally robust model predictive control approach towards the design of adaptive cruise control (ACC) systems. We model the preceding vehicle as an autonomous stochastic system, using a hybrid model with continuous dynamics and discrete, Markovian inputs. We estimate the (unknown) transition probabilities of this model empirically using observed mode transitions and simultaneously determine sets of probability vectors (ambiguity sets) around these estimates, that contain the true transition probabilities with high confidence. We then solve a risk-averse optimal control problem that assumes the worst-case distributions in these sets. We furthermore derive a robust terminal constraint set and use it to establish recursive feasibility of the resulting MPC scheme. We validate the theoretical results and demonstrate desirable properties of the scheme through closed-loop simulations.

Keywords: Learning and adaptation in autonomous vehicles, Intelligent driver aids, Motion control

1. INTRODUCTION

In recent decades, the usage of adaptive cruise control (ACC) systems has become widespread in the automotive research and industry, as they have demonstrated numerous benefits in terms of safety, fuel efficiency, passenger comfort, etc. The term ACC generally refers to longitudinal control systems aimed at maintaining a user-specified reference velocity, while avoiding collisions with preceding vehicles.

The recently proposed *Responsibility-Sensitive Safety* (RSS) framework (Shalev-Shwartz et al., 2017), prescribes minimal safety distances for ACC systems based on simple vehicle kinematics, which can guarantee collision avoidance under natural assumptions on boundedness of the accelerations. Furthermore, the authors define rules that prescribe how an ACC system should properly respond to violations of this safety distance. Although safe, the prescribed rules are reactive in nature, which may lead to sudden braking maneuvers, reducing passenger comfort and fuel efficiency.

By contrast, model predictive control (MPC) methods optimize a specified performance index based on the predicted evolution of the controlled system in the near future, which endows the control system with the capability to behave proactively, and adapt its actions with respect to potential future events. However, due to the involvement of human actors, there is an inherent level of uncertainty in the prediction of traffic situations. In order to explicitly account for this uncertainty, stochastic MPC has been a particularly popular approach (Bichi et al. (2010); Moser et al. (2018); McDonough et al. (2013)).

In an attempt to make accurate predictions about the future behavior of the lead vehicle, many different driver models have been proposed in the literature (see Wang et al. (2014) for a survey). A common approach is to combine continuous physics-based dynamics with a discrete (and potentially stochastic) decision model for the driver (e.g., Sadigh et al. (2014); Kiencke et al. (1999); Bichi et al. (2010)). We follow this line of reasoning and model the preceding vehicle using double integrator dynamics, where the driver's inputs are generated by a Markov chain.

A major shortcoming of stochastic MPC approaches is their dependence on accurate knowledge of all probability distributions involved in the stochastic model. Since, in practice, these are estimated based on finitely sized data samples, they may not accurately reflect the true underlying distributions — we will refer to this uncertainty on probability distributions as *ambiguity*. Due to this

⋆ This work was supported by the Ford-KU Leuven Research Alliance. The work of P. Patrinos was supported by: FWO projects: No. G086318N; No. G086518N; Fonds de la Recherche Scientifique – FNRS, the Fonds Wetenschappelijk Onderzoek – Vlaanderen under EOS Project No. 30468160 (SeLMA), Research Council KU Leuven C1 project No. C14/18/068.

ambiguity, stochastic controllers may perform unreliably with respect to the true distributions.

The main contributions of this paper are to address these issues in the following manner. First, we generalize the stochastic MPC methodology for ACC systems by adopting a distributionally robust approach, where not only the estimated distribution is taken into consideration, but all distributions that belong to a so-called *ambiguity set*. Under the Markovian assumption, we can use concentration inequalities to obtain closed-form expressions for these sets, such that they contain the data-generating distributions with arbitrarily high confidence. Secondly, we derive a robust control invariant set which can be used as a terminal constraint set in the proposed control formulation, allowing us to establish guaranteed recursive feasibility of the resulting MPC scheme.

Proofs Due to space restrictions, our results are stated without proof. We refer to the extended version (Schuurmans et al., 2020) for full proofs and additional details.

1.1 Notation and preliminaries

Given two integers $a \leq b$, let $\mathbb{N}_{[a,b]} := \{n \in \mathbb{N} \mid a \leq n \leq b\}$. We define the operator $[\,\cdot\,]_+$ as $\mathbf{max}\{0, \cdot\}$, where the $\mathbf{max}$ is interpreted element-wise. We denote the element of a matrix P at row i and column j as $P_{i,j}$ and the ith row of a matrix P as P_i. Similarly, the ith element of a vector x is denoted x_i. We denote the vector in $\mathbb{R}^k$ with all elements one as $\mathbf{1}_k := (1)_{i=1}^k$. Finally, we define the indicator function as $1_{x=y} := 1$ if $x = y$ and 0 otherwise.

Risk measures and ambiguity Let Ω denote a discrete sample space endowed with the σ-algebra $\mathcal{F} = 2^\Omega$ and probability measure $\mathbb{P}$, defining the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For a given random variable $Z : \Omega \to \mathbb{R}$, we can collect the possible outcomes of Z in a *random vector* $\mathbb{R}^{|\Omega|} \ni z = (Z(i))_{i\in\Omega}$. Similarly, a *probability vector* can be defined as $\mathcal{D}_{|\Omega|} \ni \mu = (\mu_i)_{i\in\Omega} = (\mathbb{P}[\{\omega\}])_{\omega\in\Omega}$, where $\mathcal{D}_k := \{p \in \mathbb{R}^k | \mathbf{1}_k{}^\top p = 1, p \geq 0\}$ denotes the probability simplex of dimension k. A risk measure $\rho : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ is a mapping from the space of possible outcomes of Z to the real line, which we may use to deterministically compare random variables before their outcome is revealed. In particular, we are interested in so-called *coherent* risk measures, for which the following dual representation exists (Shapiro et al., 2009, Thm 6.5)

$$\rho[z] = \max_{\mu\in\mathcal{A}} \mathbb{E}^\mu[z]. \tag{1}$$

Here, $\mathcal{A} \subseteq \mathcal{D}_{|\Omega|}$ is some closed, convex subset of the probability simplex, commonly referred to as the *ambiguity set* of ρ. This dual representation allows for a distributionally robust interpretation where, based on a set of data drawn from an unknown distribution, the ambiguity set is typically constructed such that it contains all probability distributions that are in some sense consistent with the data. We will use this perspective explicitly when constructing a data-driven MPC scheme in Section 3. For a given ambiguity set $\mathcal{A}$, we will denote the induced risk measure by $\rho^{\mathcal{A}}$. We finally remark that the concept of a risk measure can be extended in a straightforward manner to *conditional risk mappings* by replacing the expectation in (1) with a conditional expectation.

2. NOMINAL STOCHASTIC MPC

In this section, we construct a model for the ACC system and formulate a nominal control problem for the simplified case where all involved probability distributions are known. We use this setting to derive a terminal constraint set that allows us to ensure recursive feasibility of the MPC scheme. In Section 3, we will extend these results to the setting in which all distributions are to be estimated from data.

2.1 Modeling and problem statement

Throughout this paper, we will assume that the behavior of the vehicle pair can be modelled as a discrete-time Markov jump linear system (MJLS) (Costa et al. (2006)), which has dynamics of the form

$$x_{t+1} = f(x_t, u_t, w_{t+1}) = A(w_{t+1})x_t + B(w_{t+1})u_t + p(w_{t+1}), \tag{2}$$

where $x_t \in \mathbb{R}^{n_x}$ is the state vector $u_t \in \mathbb{R}^{n_u}$ is the input and $\mathbf{w} := (w_t)_{t\in\mathbb{N}}$ is a Markov chain on $(\Omega, \mathcal{F}, \mathbb{P})$ with state space $\mathcal{W} := \mathbb{N}_{[1,M]}$ and transition matrix $P \in \mathbb{R}^{M\times M}$, where $P_{i,j} = \mathbb{P}[w_t = j \mid w_{t-1} = i]$. We assume that at any time t, both x_t and w_t are observable.

The goal is to select a state feedback law $\kappa : \mathbb{R}^{n_x} \times \mathcal{W} \to \mathbb{R}^{n_u}$, such that for all $t \in \mathbb{N}$, $\kappa(x_t, w_t) \in \mathcal{U}$, and that for the closed-loop system $x_{t+1} = f(x_t, \kappa(x_t, w_t), w_{t+1})$, the state satisfies

$$x_t \in \mathcal{X}_{\mathrm{r}}, \tag{3a}$$

$$\mathbb{P}[x_{t+1} \in \mathcal{X}_{\mathrm{c}} \mid x_t, w_t] \geq 1 - \delta, \tag{3b}$$

almost surely (a.s.), i.e., for all $(w_i)_{i=0}^t \in \mathcal{W}^{t+1}$ such that $P_{w_t,i} > 0$. Here, the set $\mathcal{U}$, $\mathcal{X}_{\mathrm{r}}$ and $\mathcal{X}_{\mathrm{c}}$ correspond respectively to the input constraints, hard state constraints, and soft (probabilistic) state constraints, specified below.

Dynamics We model the longitudinal dynamics of the two vehicles along a road-aligned coordinate system and combine the states of the ego vehicle and the target vehicle into one system. We denote by p_{EV} and p_{TV} the positions of the ego vehicle and the target vehicle respectively and define $h := p_{\mathrm{TV}} - p_{\mathrm{EV}}$ to be the (positive) headway between the two vehicles. Similarly, we denote the velocities of the ego and target vehicle by v_{EV} and v_{TV}, so that the total state of the vehicle pair is described by a state vector $x = \left[\, h \; v_{\mathrm{EV}} \; v_{\mathrm{TV}} \,\right]^\top$. For simplicity, we take the individual vehicle dynamics to be described by discrete double integrators, such that the combined dynamics is given by

$$x_{t+1} = \begin{bmatrix} 1 & -T_{\mathrm{s}} & T_{\mathrm{s}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ T_{\mathrm{s}} \\ 0 \end{bmatrix} u_t + \begin{bmatrix} 0 \\ 0 \\ T_{\mathrm{s}} a_{\mathrm{TV}}(x_t, u_t, w_{t+1}) \end{bmatrix}, \tag{4}$$

where T_{s} is the sampling period and a_{TV} denotes a mode-dependent acceleration of the target vehicle. Provided that a_{TV} is an affine function of the states and inputs, this model is compatible with the form (2). In the remainder of this paper, we assume a parametrization of a_{TV} such that in decelerating modes, the input (the brake) behaves like a dissipative element, i.e.,

$$a_{\mathrm{TV}}(x, u, w) = a_{\mathrm{TV}}(x, w) = \begin{cases} c_w, & \text{if } c_w \geq 0 \\ c_w x_3, & \text{otherwise}, \end{cases} \tag{5}$$

where $c_w \geq -1/T_{\mathrm{s}}, w \in \mathcal{W}$ are design parameters.

Constraints We assume that velocities of the ego vehicle must remain nonnegative and upper bounded by some physical limit $v_{\max} > 0$, and that the acceleration of the target vehicle is limited between the values $a_{\min} \leq 0$ and $a_{\max} \geq 0$. This yields the constraint sets

$$\mathcal{X}_{\mathrm{r}} := \{x \in \mathbb{R}^{n_x} \mid 0 \leq x_2 \leq v_{\max}\}, \tag{6a}$$
$$\mathcal{U} := \{u \in \mathbb{R}^{n_u} \mid a_{\min} \leq u \leq a_{\max}\}, \tag{6b}$$

for the states and inputs respectively. We do not pose explicit constraints on the state x_3 as we assume that the controller has no agency over the target vehicle.

Since a stochastic model of the target vehicle will typically include extreme behaviors, albeit with exceedingly small probabilities, imposing certain safety constraints robustly (i.e., for all possible realizations of $\mathbf{w}$) will often lead to overly large safety distances, excessive emergency maneuvers, or even infeasibility of the optimization problem in practically benign situations. It is therefore common to instead impose (conditional) chance constraints of the form (3b) (e.g., Moser et al. (2018)). In particular, we want to constrain the headway (possibly defined to include some safety distance), to remain positive:

$$\mathcal{X}_{\mathrm{c}} = \{x \in \mathbb{R}^{n_x} \mid g(x) = -x_1 \leq 0\}.$$

Since chance constraints (3b) are generally nonconvex, it is common to approximate them using risk measures (Nemirovski (2012)). In particular, it can be shown (Shapiro et al., 2009, sec. 6.2.4) that for any random variable $z \sim p \in \mathcal{D}_m$, the following implication holds tightly

$$\mathrm{AV@R}_\delta^p[z] \leq 0 \Rightarrow \mathbb{P}[z \leq 0] \geq 1 - \delta. \tag{7}$$

Here, $\mathrm{AV@R}_\delta^p[z]$ denotes a particular risk measure referred to as the *average value-at-risk* (at level $\delta \in (0, 1]$ and with reference probability $p \in \mathcal{D}_m$). It can be defined through the form (1), using its polytopic ambiguity set Sopasakis et al. (2019)

$$\mathcal{A} = \mathcal{A}_{\mathrm{AV@R}_\delta^p} := \{\mu \in \mathbb{R}^{|\Omega|} \mid \mathbf{1}_{|\Omega|}{}^\top \mu = 1, 0 \leq \mu \leq \tfrac{p}{\delta}\}. \tag{8}$$

By exploiting the structure of ambiguity sets such as $\mathcal{A}_{\mathrm{AV@R}_\delta^p}$, Sopasakis et al. (2019) show that constraints involving the average value-at-risk can be imposed efficiently using only linear (in)equalities. We can thus satisfy the chance constraint (3b) by imposing for $t \in \mathbb{N}$,

$$\mathrm{AV@R}_\delta^{P_{w_t}}[g(x_{t+1}) \mid x_t, w_t] \leq 0, \text{ a.s.} \tag{9}$$

Finally, in order to guarantee recursive feasibility, we impose the final state to be in a robust control invariant set $x_N \in \mathcal{X}_N$ for all $(w_i \in \mathcal{W})_{i \in \mathbb{N}_{[0,N]}}$. This set is specified in Section 2.2.

Cost function We define a stage cost $\ell : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}_+$ and terminal cost $\ell_N : \mathbb{R}^{n_x} \to \mathbb{R}_+$, that simply assign a quadratic penalty to the deviation from the reference velocity v_{ref} and to the control effort u:

$$\ell(x, u) := q(x_2 - v_{\mathrm{ref}})^2 + ru^2, \quad \ell_N(x) := q(x_2 - v_{\mathrm{ref}})^2.$$

Definition 1. (Nominal stochastic MPC). For a given $x \in \mathbb{R}^{n_x}, w \in \mathcal{W}$, the nominal optimal control problem (OCP) comprises of computing an N-step sequence of admissible policies, i.e., a sequence of functions $\pi = (\kappa_i)_{i \in \mathbb{N}_{[0,N-1]}}$, with $\kappa_k : \mathbb{R}^{n_x} \times \mathcal{W} \to \mathbb{R}^{n_u}$ that solve the optimization problem

$$\begin{aligned}\underset{u_0}{\mathbf{minimize}}\ & \ell(x_0, u_0) + \inf_{u_1} \mathbb{E}_{|0}\Big[\ell(x_1, u_1) + \dots \\ & + \inf_{u_{N-1}} \mathbb{E}_{|N-2}\big[\ell(x_{N-1}, u_{N-1}) + \mathbb{E}_{|N-1}\big[\ell_N(x_N)\big] \cdots\big]\Big]\end{aligned} \tag{10a}$$

subject to

$$x_0 = x, w_0 = w, \tag{10b}$$
$$x_{k+1} = f(x_k, u_k, w_{k+1}),\ k \in \mathbb{N}_{[0,N-1]}, \tag{10c}$$
$$u_k = \kappa_k(x_k, w_k) \in \mathcal{U}, x_k \in \mathcal{X}_{\mathrm{r}} \text{ a.s.},\ k \in \mathbb{N}_{[0,N-1]}, \tag{10d}$$
$$\mathrm{AV@R}_\delta^{P_{w_k}}[g(x_{k+1}) \mid x_k, w_k] \leq 0 \text{ a.s.}, k \in \mathbb{N}_{[0,N-1]}, \tag{10e}$$
$$x_N \in \mathcal{X}_N \text{ a.s.}, \tag{10f}$$

where $\mathbb{E}_{|t}[\,\cdot\,] = \mathbb{E}[\,\cdot\,|x_t, w_t]$ denotes the conditional expectation given the realization of $(w_i)_{i=0}^t$.

The corresponding MPC scheme is obtained by applying the first policy κ_0 to the system at the current state, and resolving the OCP (10) in a receding horizon manner. Due to the discrete nature of $\mathcal{W}$, problem (10) can be stated as a finite-dimensional optimization problem over a *scenario tree* (Sopasakis et al. (2019)).

Remark 2. Note that by linearity of the expectation operator, the cost (10a) is equivalent to the total expectation of the sum of the state costs $\ell(x_t, u_t)$ and the terminal cost $\ell_N(x_N)$. However, by writing the cost in the nested form above, we emphasize the relation with the risk-averse OCP formulated in Section 3.

2.2 Recursive feasibility of the nominal problem

In this section, we describe a simple procedure to obtain a robust control invariant set $\mathcal{X}_N$ and use it to establish recursive feasiblity of the nominal stochastic MPC scheme.

Definition 3. (Robust control invariant set). Let $\mathcal{X}$ denote a set of feasible states and $\mathcal{U}$ the set of feasible control actions. A set $\mathcal{R} \subseteq \mathcal{X}$ is called a *robust control invariant* (RCI) set for the system (2) if for all $x \in \mathcal{R}$, there exists a $u \in \mathcal{U}$ such that $f(x, u, w) \in \mathcal{R}, \forall w \in \mathcal{W}$.

Definition 4. (Maximal robust control invariant set). An RCI set $\mathcal{R}^\star$ is called the *maximal robust control invariant* (MRCI) set if $\mathcal{R} \subseteq \mathcal{R}^\star$ for all RCI sets $\mathcal{R}$.

Definition 5. (Robust positively invariant set). A set $\mathcal{R}_\kappa \subseteq \mathcal{X}$ is a *robust positively invariant* (RPI) set for the system (2) driven by the control law $\kappa : \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$, if for all $x \in \mathcal{R}_\kappa$, it holds that $\kappa(x) \in \mathcal{U}$ and $f(x, \kappa(x), w) \in \mathcal{R}_\kappa, \forall w \in \mathcal{W}$. Note that any RPI set is necessarily RCI.

For notational convenience, we construct a set $\mathcal{Z}_{\mathrm{s}} \subseteq \mathcal{X}_{\mathrm{r}} \times \mathcal{W}$, akin to the *stochastic feasibility set* defined by Korda et al. (2011). It contains all augmented states (x, w) that are feasible and for which a feasible input exists, with respect to both the soft constraints (10e) and hard constraints (10d):

$$\mathcal{Z}_{\mathrm{s}} := \left\{ (x, w) \,\middle|\, \begin{array}{c} x \in \mathcal{X}_{\mathrm{r}}, w \in \mathcal{W}, \exists u \in \mathcal{U} : \\ \mathrm{AV@R}_\delta^{P_w}[g(f(x, u, w')) \mid (x, w)] \leq 0 \\ w' \sim P_w, \end{array} \right\}. \tag{11}$$

Our goal is to compute a sufficiently large terminal constraint set $\mathcal{X}_N$, such that $\mathcal{X}_N \times \mathcal{W} \subseteq \mathcal{Z}_{\mathrm{s}}$. To this end, we first explicitly define a simple polyhedral RPI subset of $\mathcal{X}_{\mathrm{r}}$ for the system (4) as shown in the following result. By iteratively expanding this set, we can then obtain an inner approximation of the MRCI set.

Let $c_{\min} := \mathbf{min}_{w\in\mathcal{W}} c_w$ denote the parameter of the target vehicle model (5) corresponding to the maximal deceleration. Recall that we assumed that $-1/T_s \leq c_{\min} < 0$.

Proposition 6. Let us define the linear state feedback policy $u = Kx$, where $K := [\,0\ c_{\min}\ 0\,]$, and the corresponding candidate RPI set

$$\mathcal{R}_K := \left\{ x \in \mathbb{R}^{n_x} \,\middle|\, \begin{array}{l} \frac{a_{\min}}{c_{\min}} \geq x_2 \geq 0,\ x_3 \geq x_2, \\ v_{\max} \geq x_2,\ g(x) = -x_1 \leq 0 \end{array} \right\}.$$

The following statements hold: (i) $\mathcal{R}_K$ is RPI for the dynamics (4) and policy $u = Kx$; (ii) $Kx \in \mathcal{U}$ for every $x \in \mathcal{R}_K$, with $\mathcal{U}$ as defined in (6b); and (iii) $\mathcal{R}_K \times \mathcal{W} \subseteq \mathcal{Z}_s$.

We can now iteratively expand $\mathcal{R}_K$, to obtain the following iterates (Kerrigan, 2000, Alg. 2.1)

$$\mathcal{R}^{(i+1)} = \mathbf{pre}\left(\mathcal{R}^{(i)}\right) \cap \mathcal{X}_r \cap \mathcal{X}_c,\ \mathcal{R}^{(0)} = \mathcal{R}_K,$$

where $\mathbf{pre}(\mathcal{R}^{(i)}) := \{x \in \mathbb{R}^{n_x} | \exists u \in \mathcal{U} : f(x,u,w) \in \mathcal{R}^{(i)}, \forall w \in \mathcal{W}\}$ denotes the pre-set of $\mathcal{R}^{(i)}$. Note that since all involved sets are polyhedral, the pre-set can be easily computed using standard techniques (Borrelli et al. (2017)). From (Kerrigan, 2000, Prop. 2.6.1), it then follows that for all $i \in \mathbb{N}$, $\mathcal{R}^{(i)}$ is RCI. Therefore, we may choose to terminate after any finite number of iterations, and still retain guaranteed recursive feasibility .

Definition 7. (Recursive feasibility). An MPC controller is recursively feasible if the existence of a feasible solution $\pi_{|t} = (\kappa_i)_{i\in\mathbb{N}_{[0,N-1]}}$ to the optimal control problem with initial state $(x,w) \in \mathcal{Z}_s$ implies *almost surely* that there exists a feasible solution to the optimal control problem with initial state $(f(x,\kappa_0(x,w),w'),w')$, $w' \sim P_w$.

Theorem 8. (Nominal recursive feasibility). If $\mathcal{X}_N$ is RCI and $\mathcal{X}_N \times \mathcal{W} \subseteq \mathcal{Z}_s$, then the nominal stochastic MPC problem is recursively feasible.

3. DISTRIBUTIONALLY ROBUST FORMULATION

We now move to the more realistic setting in which the measure $\mathbb{P}$, and by extension the transition matrix $P \in \mathbb{R}^{M\times M}$ governing the Markov chain is unknown. In this setting, we need to resort to data-driven estimates of the transition probabilities, which are subject to some level of ambiguity. Taking this into account in a distributionally robust manner leads to a modified version of the MPC problem (10).

3.1 From Markovian data to ambiguity sets

Suppose we are given a sample $W = \{w_i\}_{i=1}^n$ of n observations from the Markov chain with unknown transition matrix P. To simplify matters, we partition W into subsets $W_j \subseteq W$, $j \in \mathcal{W}$, which contain only the transitions that originated in mode j. That is, $W_j := \{w_i \in \mathcal{W} \mid w_{i-1} = j\}$. Due to the Markov property, the samples $w \in W_j$ are independent and identically distributed (i.i.d.) with distribution P_j, i.e., the jth row of the transition matrix. We compute the empirical distributions of W_j to obtain estimates $\hat{P}_j$ for the transition probabilities. That is,

$$\hat{P}_{j,i} := \begin{cases} \frac{1}{n_j}\sum_{w\in W_j} 1_{w=i}, & \text{if } n_j > 0, \\ \frac{1}{M}, & \text{otherwise,} \end{cases} \tag{12}$$

for all $i,j \in \mathcal{W}$, where $n_j := |W_j|$ is the number of samples in each subset of the data. Given an arbitrary confidence level $\alpha \in (0,1]$, we can now for each such estimate $\hat{P}_j$, use the results in Schuurmans et al. (2019) to define an *ambiguity set*

$$\mathcal{A}_{r_j}^{\ell_1}(\hat{P}_j) := \{p \in \mathcal{D}_M \mid \|p - \hat{P}_j\|_1 \leq r_j\}, \tag{13}$$

where r_j is computed such that $\mathbb{P}[P_j \in \mathcal{A}_{r_j}^{\ell_1}(\hat{P}_j)] \geq 1 - \alpha$.

By the dual risk representation (1), the computed ambiguity sets $\mathcal{A}_{r_j}^{\ell_1}(\hat{P}_j)$ implicitly define coherent risk measures. Thus, by replacing the now unknown probability distributions in (10) by the worst-case distributions in the ambiguity sets, we transform it to a *risk-averse* MPC problem (Sopasakis et al. (2019)), in which the ambiguity in the estimated transition matrices is accounted for.

By collecting additional data samples during closed-loop operation (i.e., by increasing n_j and therefore decreasing r_j for all modes j) the ambiguity sets will asymptotically shrink to the singletons $\{P_j\}$. As such, conservatism of the controller is gradually reduced while constraint satisfaction with respect to the true distributions is guaranteed.

3.2 Risk-averse MPC formulation

Cost function The proposed distributionally robust approach replaces the conditional expectations by conditional risk mappings based on the risk measures induced by the ambiguity sets (13). For ease of notation we will for a given sequence of ambiguity sets $\bar{\mathcal{A}} := (\mathcal{A}_j)_{j\in\mathcal{W}}$, denote the conditional risk mapping of the random stage costs as

$$\rho_{|t}^{\bar{\mathcal{A}}}[\ell(x_{t+1},u_{t+1})] := \max_{p\in\mathcal{A}_{w_t}} \mathbb{E}_{\delta}^{p}[\ell(x_{t+1},u_{t+1}) \mid x_t, w_t].$$

Ambiguous chance constraints Since the implication (7) holds only with respect to the true but unknown probability measure $\mathbb{P}$, the risk constraint (10e) no longer guarantees satisfaction of the original chance constraints in the current setting. We will therefore impose it robustly with respect to all distributions in the data-driven ambiguity sets $\mathcal{A}_{r_j}^{\ell_1}(\hat{P}_j)$, leading to the following definition.

Definition 9. (Distributionally robust AV@R). Given a random vector $z \in \mathbb{R}^n$ and an ambiguity set $\mathcal{A} \subseteq \mathcal{D}_n$, we define the distributionally robust average value-at-risk of z as

$$\mathbf{r}\text{-AV@R}_{\delta}^{\mathcal{A}}[z] := \max_{\nu\in\mathcal{A}} \text{AV@R}_{\delta}^{\nu}[z]. \tag{14}$$

For the ℓ_1-ambiguity set $\mathcal{A} = \mathcal{A}_r^{\ell_1}(\hat{p})$ of radius r around an empirical estimate $\hat{p}$, we can use the definitions (13) and (8) of $\mathcal{A}_r^{\ell_1}(\hat{p})$ and $\mathcal{A}_{\text{AV@R}_{\delta}^{\nu}}$ to express (14) explicitly as

$$\mathbf{r}\text{-AV@R}_{\delta}^{\mathcal{A}_r^{\ell_1}(\hat{p})}[z] = \max_{\pi,\nu\in\mathcal{D}_n} \{\pi^\top z \mid \|\nu - \hat{p}\|_1 \leq r, \pi \leq \nu/\delta\}.$$

Recall that we assume that the radius r in the definition of the ambiguity set is chosen to satisfy $\mathbb{P}[p \in \mathcal{A}_r^{\ell_1}(\hat{p})] \geq 1-\alpha$. Therefore we have that with probability at least $(1-\alpha)$, $\text{AV@R}_{\delta}^{p}[z] \leq \mathbf{r}\text{-AV@R}_{\delta,r}^{\mathcal{A}_r^{\ell_1}(\hat{p})}[z]$, so that a constraint on a random value z of the form $\mathbf{r}\text{-AV@R}_{\delta,r}^{\mathcal{A}_r^{\ell_1}(\hat{p})}[z] \leq 0$, implies that $\mathbb{P}[z \leq 0] \geq 1-\epsilon$, where $1-\epsilon \geq (1-\delta)(1-\alpha)$. Thus, by replacing the AV@R risk measure used in the conditional risk constraints (10e) by $\mathbf{r}$-AV@R, satisfaction of chance constraint can still be guaranteed despite the incomplete knowledge of the transition matrix. We summarize these modifications in the following definition.

Definition 10. (Risk-averse MPC problem). For a given initial state $x \in \mathbb{R}^{n_x}, w \in \mathcal{W}$, and sequence of ambiguity sets $\bar{\mathcal{A}} := (\mathcal{A}_j \subseteq \mathcal{D}_M)_{j\in\mathcal{W}}$, the risk-averse OCP comprises of computing an N-step sequence of admissible policies $\pi = (\kappa_i)_{i\in\mathbb{N}_{[0,N-1]}}$, with $\kappa_k : \mathbb{R}^{n_x} \times \mathcal{W} \to \mathbb{R}^{n_u}$ that solve the optimization problem

$$\underset{u_0}{\textbf{minimize}}\ \ell(x_0,u_0) + \inf_{u_1} \rho^{\bar{\mathcal{A}}}_{|0}\Big[\ell(x_1,u_1) + \dots + \inf_{u_{N-1}} \rho^{\bar{\mathcal{A}}}_{|N-2}\big[\ell(x_{N-1},u_{N-1}) + \rho^{\bar{\mathcal{A}}}_{|N-1}\big[\ell_N(x_N)\big]\cdots\big]\Big] \tag{15a}$$

subject to

$$x_0 = x, w_0 = w, \tag{15b}$$
$$x_{k+1} = f(x_k, u_k, w_{k+1}), \tag{15c}$$
$$u_k = \kappa_k(x_k, w_k) \in \mathcal{U}, x_k \in \mathcal{X}_{\mathrm{r}}, \tag{15d}$$
$$\textbf{r-}\mathrm{AV@R}^{\mathcal{A}_{w_k}}_{\delta}[g(x_{k+1}) \mid x_k, w_k] \le 0, \tag{15e}$$
$$x_N \in \mathcal{X}_N, \forall \mathbf{w}_N \in \mathcal{W}^N, \tag{15f}$$

$\forall \mathbf{w}_k \in \mathcal{W}^k, \forall k \in \mathbb{N}_{[0,N-1]}$, where we introduced the shorthand $\mathbf{w}_k := (w_i)_{i=1}^k$.

Remark 11. Without knowledge of the true distributions, imposing constraints *almost surely* – even for all distributions in the ambiguity set – is no longer sufficient to guarantee recursive feasibility, since with a probability of at most $\alpha > 0$, a nonzero transition probability to a given mode may not be reflected in any probability vector in the used ambiguity set. In this case, a feasible solution at a given time cannot be used to guarantee the existence of a feasible solution in the next. Therefore, we impose constraints at stage k *for all* realizations of $\mathbf{w}_k$.

Theorem 12. (Risk-averse recursive feasibility). If for all time steps t and $t+1$, the risk-averse MPC problem (15) is instantiated with ambiguity sets $\bar{\mathcal{A}}_t = (\mathcal{A}_{t,j})_{j\in\mathcal{W}}$ and $\bar{\mathcal{A}}_{t+1} = (\mathcal{A}_{t+1,j})_{j\in\mathcal{W}}$, such that $\mathcal{A}_{t+1,j} \subseteq \mathcal{A}_{t,j}, \forall j \in \mathcal{W}$, then, the MPC scheme is recursively feasible.

4. NUMERICAL SIMULATIONS

Terminal constraint sets For the considered set-up, the RSS model described in Shalev-Shwartz et al. (2017) derives a minimal safety distance required for guaranteed collision avoidance. It involves computing the distances $\Delta_{\mathrm{EV}}(x_2), \Delta_{\mathrm{TV}}(x_3)$ required for the ego vehicle and target vehicle respectively to come to a halt in an emergency braking scenario, as a function of their initial velocities x_2, x_3. The minimal required distance is given by $h_{\min,\mathrm{RSS}}(x_2,x_3) := [\Delta_{\mathrm{EV}}(x_2) - \Delta_{\mathrm{TV}}(x_3)]_+$. Although derived for continuous-time systems, the derivation can be easily repeated for the discrete-time model at hand. It has to be noted, however, that in general, $\mathcal{X}_{\mathrm{RSS}} := \{x \mid x_1 \ge h_{\min,\mathrm{RSS}}(x_2,x_3)\}$ is not RCI for the system (4). Similarly, for a given pair of velocities x_2 and x_3, the iteratively computed terminal constraint sets $\mathcal{R}^{(i)}$ can be associated to a minimal safety distance $h^{(i)}_{\min}(x_2,x_3) := \mathbf{min}\{h \mid [\,h\ x_2\ x_3\,]^\top \in \mathcal{R}^{(i)}\}$, where we set $h^{(i)}_{\min} = \infty$ if no feasible solution exists. Figure 1 shows the safety distance according to both approaches as a function of x_2. Note that the initial set $\mathcal{R}^{(0)}$ is more conservative than RSS. However, after $i = 12$ iterations, $\mathcal{R}^{(i)}$ has converged and yields a smaller safety distance than RSS for all values of x_2. Thus, we find that in practice, the requirement of the terminal set to be RCI introduces no conservatism over the hand-crafted safety distance provided by RSS.

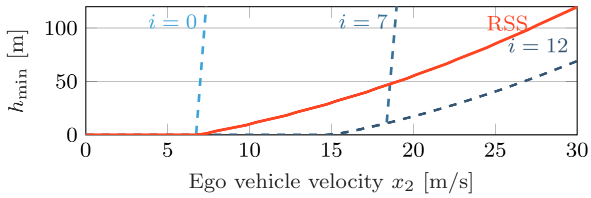

Figure 1. Minimal safety distances $h^{(i)}_{\min}$ and $h_{\min,\mathrm{RSS}}$, for $v_{\max} = 40\mathrm{m/s}$, $a_{\min} = -5\mathrm{m/s}^2$, $c_{\min} = -0.33\mathrm{s}^{-1}$ and a fixed target vehicle velocity $x_3 = 20\mathrm{m/s}$.

Closed-loop simulations The following experiments demonstrate the benefit of the proposed learning-based MPC scheme in Section 3 (referred to as the *risk-averse* approach), as compared to the two extreme variants obtained by taking $\mathcal{A}_j = \{\hat{P}_j\}$ and $\mathcal{A}_j = \mathcal{D}_M$, $j \in \mathcal{W}$ (referred to as the *stochastic* and *robust* approach, respectively). For the stochastic approach, we set the tolerated chance constraint violation probability to $\delta_{\mathrm{s}} = 0.1$, and for the risk-averse controller, we choose $\alpha = \delta = 0.05$, such that $(1-\alpha)(1-\delta) \approx 1 - \delta_{\mathrm{s}}$. All used controller settings are as summarized in Table 1, unless otherwise specified. The (unknown) transition matrices used in the experiments are

$$P_p = \begin{bmatrix} 0.92 & 0.04 & 0.02 & 0.02 \\ 0.29 & 0.50 & 0.09 & 0.12 \\ 0.26 & 0.21 & 0.36 & 0.17 \\ 0.31 & 0.25 & 0.23 & 0.21 \end{bmatrix} \text{ and } P_s = \begin{bmatrix} 0.29 & 0.7 & 0.009 & 0.001 \\ 0.09 & 0.90 & 0.009 & 0.001 \\ 0.4 & 0.29 & 0.3 & 0.01 \\ 0.048 & 0.001 & 0.001 & 0.95 \end{bmatrix}.$$

The optimal control problems are formulated using Yalmip (Löfberg (2004)) and solved using MOSEK (MOSEK ApS (2017)) on an Intel Core i7-7700K CPU.

Table 1. Default controller settings.

(q,r)	T_{s} [s]	N	$(v_{\mathrm{ref}}, v_{\max})$ [m/s]	$(a_{\min}, a_{\max})$ [m/s^2]
(5, 10)	0.5	3	(30, 40)	(−4, 5)

Performance For a fixed initial state, we performed 100 randomized simulations of 50 time steps for the three controllers with prediction horizon $N = 5$. The target vehicle parameters are $(c_i)_{i\in\mathcal{W}} = [\,1.13\ -0.02\ -0.33\ -0.16\,]$ and the true transition matrix is set to $P = P_p$. The average solver time for these experiments was 0.45s. We compare the performance of the controllers by computing the closed-loop cost over each realization. We conducted this experiment both without and with offline learning. In the former case, all transition probabilities are estimated online, whereas in the latter, a sequence of 5000 draws from the Markov chain is provided to the controller before deployment. Figure 2 shows the empirical cumulative distribution of the closed-loop costs with and without offline learning. We observe that due to the initial lack of data, the risk-averse controller selects a large ambiguity set, which renders its behavior indistinguishable from that of the robust controller. The stochastic approach, on the other hand, introduces no such conservatism and thus achieves lower costs more frequently than the competing controllers. As the risk-averse controller observes more data (Figure 2, right), its conservatism decreases, allowing it to achieve a cost distribution that closely resembles that of the stochastic approach, while still providing the same recursive feasibility guarantees as the robust approach.

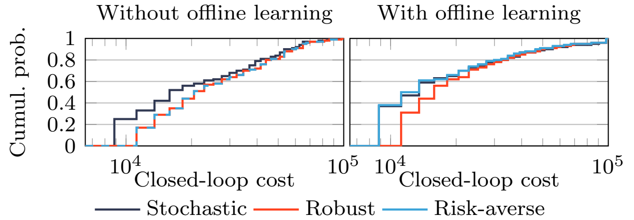

Figure 2. Empirical cumulative distribution of the closed-loop cost over 100 randomized simulations.

Safety In the following experiment, we use the target vehicle parameters $(c_i)_{i \in \mathcal{W}} = [\,1.1\ 0\ -0.5\ -1\,]$ and transition matrix $P = P_s$. We simulate a low-probability emergency situation by forcing the Markov chain to switch to mode 4 at a single fixed time step during each simulation, corresponding to a harsh braking maneuver of the target vehicle. Note that from any mode $i \in \mathcal{W}$, there is a nonzero switching probability to mode 4. Therefore, the simulated trajectories correspond to possible realizations for which infeasibility of the OCP is not acceptable. We repeated this simulation for 100 realizations of 200 steps, and with increasing sample sizes n for offline learning. The average solver time for this experiment was 0.036s. Figure 3 shows that with minimal offline learning, the stochastic controller fails to find a feasible solution in 38% of realizations. As n increases and estimated distributions become more accurate, this fraction decreases, yet it requires a sample size $n = 5000$ to reduce the number of infeasible realizations to zero for this particular experiment. By contrast, Theorem 12 guarantees recursive feasibility for the risk-averse and the robust approach regardless of n, as confirmed by the experiment.

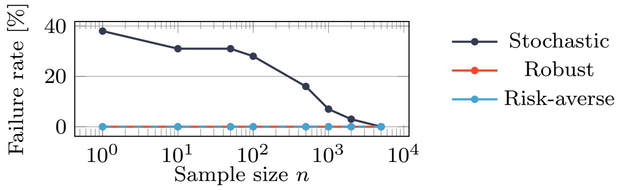

Figure 3. Percentage of infeasible realizations for the emergency braking scenario (out of 100 realizations).

5. CONCLUSION

We proposed a learning risk-averse approach towards MPC for ACC applications with Markovian driver models. This framework allows us to use collected data to improve performance of the controller with respect to the robust approach, while retaining safety guarantees through provable recursive feasibility. These benefits were illustrated through simulations. In future work, we plan to extend the methodology to more general problem set-ups and perform more extensive experiments using real-world driving data.

REFERENCES

Bichi, M., Ripaccioli, G., Cairano, S.D., Bernardini, D., Bemporad, A., and Kolmanovsky, I.V. (2010). Stochastic model predictive control with driver behavior learning for improved powertrain control. In *49th IEEE Conference on Decision and Control (CDC)*, 6077–6082.

Borrelli, F., Bemporad, A., and Morari, M. (2017). *Predictive control for linear and hybrid systems*. Cambridge University Press, Cambridge, United Kingdom; New York, NY, USA.

Costa, O.L.V., Fragoso, M.D., and Marques, R.P. (2006). *Discrete-time Markov jump linear systems*. Springer Science & Business Media.

Kerrigan, E.C. (2000). *Robust Constraint Satisfaction: Invariant Sets and Predictive Control*. Ph.D. thesis.

Kiencke, U., Majjad, R., and Kramer, S. (1999). Modeling and performance analysis of a hybrid driver model. *Control Engineering Practice*, 7(8), 985–991.

Korda, M., Gondhalekar, R., Cigler, J., and Oldewurtel, F. (2011). Strongly feasible stochastic model predictive control. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, 1245–1251. IEEE.

Löfberg, J. (2004). Yalmip : A toolbox for modeling and optimization in matlab. In *Proceedings of the CACSD Conference*. Taipei, Taiwan.

McDonough, K., Kolmanovsky, I., Filev, D., Yanakiev, D., Szwabowski, S., and Michelini, J. (2013). Stochastic dynamic programming control policies for fuel efficient vehicle following. In *2013 American Control Conference*, 1350–1355.

MOSEK ApS (2017). *The MOSEK optimization toolbox for MATLAB manual. Version 8.1.* URL `http://docs.mosek.com/8.1/toolbox/index.html`.

Moser, D., Schmied, R., Waschl, H., and Re, L.d. (2018). Flexible Spacing Adaptive Cruise Control Using Stochastic Model Predictive Control. *IEEE Transactions on Control Systems Technology*, 26(1), 114–127.

Nemirovski, A. (2012). On safe tractable approximations of chance constraints. *European Journal of Operational Research*, 219(3), 707–718.

Sadigh, D., Driggs-Campbell, K., Puggelli, A., Li, W., Shia, V., Bajcsy, R., Sangiovanni-Vincentelli, A., Sastry, S.S., and Seshia, S. (2014). Data-driven probabilistic modeling and verification of human driver behavior. In *2014 AAAI Spring Symposium Series*.

Schuurmans, M., Katriniok, A., Tseng, H.E., and Patrinos, P. (2020). Learning-Based Risk-Averse Model Predictive Control for Adaptive Cruise Control with Stochastic Driver Models. *arXiv:2005.02646 [cs, eess]*.

Schuurmans, M., Sopasakis, P., and Patrinos, P. (2019). Safe learning-based control of stochastic jump linear systems: a distributionally robust approach. *arXiv preprint arXiv:1903.10040*.

Shalev-Shwartz, S., Shammah, S., and Shashua, A. (2017). On a Formal Model of Safe and Scalable Self-driving Cars. *arXiv:1708.06374*.

Shapiro, A., Dentcheva, D., and Ruszczyński, A. (2009). *Lectures on stochastic programming: modeling and theory*. SIAM.

Sopasakis, P., Schuurmans, M., and Patrinos, P. (2019). Risk-averse risk-constrained optimal control. In *2019 18th European Control Conference (ECC)*, 375–380.

Sopasakis, P., Herceg, D., Bemporad, A., and Patrinos, P. (2019). Risk-averse model predictive control. *Automatica*, 100, 281–288.

Wang, W., Xi, J., and Chen, H. (2014). Modeling and recognizing driver behavior based on driving data: A survey. *Mathematical Problems in Engineering*, 2014.