ARTICLES
https://doi.org/10.1038/s42256-020-0225-y

nature machine intelligence

Using online verification to prevent autonomous vehicles from causing accidents

Christian Pek[1,2], Stefanie Manzinger[1,2], Markus Koschi[1,2] and Matthias Althoff[1]

Ensuring that autonomous vehicles do not cause accidents remains a challenge. We present a formal verification technique for guaranteeing legal safety in arbitrary urban traffic situations. Legal safety means that autonomous vehicles never cause accidents although other traffic participants are allowed to perform any behaviour in accordance with traffic rules. Our technique serves as a safety layer for existing motion planning frameworks that provide intended trajectories for autonomous vehicles. We verify whether intended trajectories comply with legal safety and provide fallback solutions in safety-critical situations. The benefits of our verification technique are demonstrated in critical urban scenarios, which have been recorded in real traffic. The autonomous vehicle executed only safe trajectories, even when using an intended trajectory planner that was not aware of other traffic participants. Our results indicate that our online verification technique can drastically reduce the number of traffic accidents.

Safety remains a major challenge in the realization of autonomous vehicles. Unsafe decisions by autonomous vehicles can endanger human lives and cause tremendous economic loss in terms of product liability. Although autonomous driving is becoming a reality, recent accidents involving autonomous driving systems have raised major concerns in various institutions[1], and policy makers continue to debate about adequate safety levels for certifying autonomous vehicles[2]. To achieve widespread acceptance, safety concerns must be resolved to the full satisfaction of all road users. So far, automotive safety relies primarily on simulation and testing. However, due to the infinitely many unique real-world scenarios, these techniques cannot ensure strict safety levels[3,4], especially when using machine learning for motion planning[5].

We call for a paradigm shift from accepting residual collision risks to ensuring safety through formal verification. Formal verification describes the process of proving that a system always fulfils a desired formal specification[6]. However, in the context of safe motion planning, specifying all unsafe scenarios and proper reactions of autonomous vehicles is a tedious task[6]. Although it cannot be excluded that autonomous vehicles are involved in accidents, such as when a following car deliberately provokes a rear-end collision, self-inflicted accidents can and should be eliminated. What can we expect from human drivers to avoid self-inflicted accidents? Based on the Vienna Convention on Road Traffic, which serves as a foundation for safe driving in 78 countries, human drivers 'shall avoid any behaviour likely to endanger or obstruct traffic' (article 7 of ref. [7]). Inspired by this general rule, we demand that motions of autonomous vehicles must be collision-free under the premise that other traffic participants are allowed to perform all legal behaviours, that is, all dynamically feasible behaviours that do not violate traffic rules. Following refs. [8,9], we refer to this specification as 'legal safety'.

In contrast to related work, our holistic approach computes all legal behaviours of other traffic participants and collision-free fallback plans for the autonomous vehicle. Our solution serves as a safety layer for existing motion planning frameworks. These frameworks generate intended trajectories but cannot guarantee legal safety. However, in combination with our verification technique, legal safety is ensured. Our technique provides the following three key features:

1. Online situation assessment: The safety of each traffic situation is assessed online during operation of the autonomous vehicle by rigorously predicting all legal future evolutions of the scenario (blue areas, Fig. 1) while accounting for measurement uncertainties. In contrast to classical testing approaches, even previously unseen traffic environments can be handled, that is, scenarios with arbitrary road geometries and number of traffic participants.
2. Fail-safe operation: Our approach ensures that the autonomous vehicle always has a fail-safe trajectory to a standstill in designated safe areas, which serves as a fallback plan in the case where a safety-critical situation occurs (see the fail-safe trajectory in Fig. 1).
3. Correct by construction: Regardless of the provided motion planning framework, which may include machine learning components, our verification technique ensures that the autonomous vehicle operates in compliance with legal safety at all times. Furthermore, our safety guarantees hold even if certain traffic rules are not yet included in our technique, because, from the set of all dynamically feasible behaviours, we only remove the behaviours that are illegal according to the considered traffic rules.

At present, verification is performed during the design process—that is, offline, before systems are deployed[10]. However, offline verification is not suitable for autonomous vehicles, as these vehicles operate in highly uncertain environments in which each scenario is unique. For this reason, online verification approaches have been introduced that verify safety properties during operation of the autonomous vehicles (section II-C of ref. [11]), for example, through logical reasoning[12,13] or avoiding inevitable collision states[14,15]. In the case where a trajectory is classified as unsafe, these approaches usually do not provide an alternative safe plan for the vehicle. In the field of control, popular safety techniques are robust model predictive control approaches[16–18] and correct-by-construction controllers,

[1] Cyber-Physical Systems Group, Department of Informatics, Technical University of Munich, Garching, Germany. [2] These authors contributed equally. Christian Pek, Stefanie Manzinger, Markus Koschi. e-mail: christian.pek@tum.de; stefanie.manzinger@tum.de; markus.koschi@tum.de

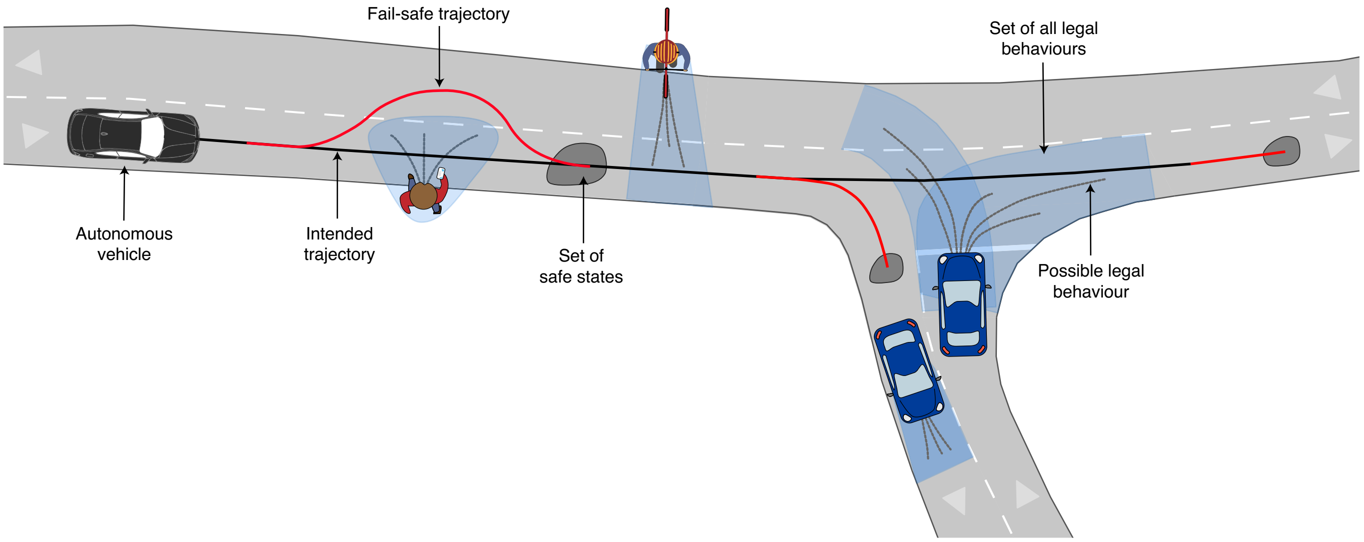

Fig. 1 | Verification of legal safety. Intended trajectories (black line) are usually planned by only considering the most likely behaviours (grey lines) of other traffic participants. Our online verification technique ensures that the autonomous vehicle is safe in accordance with legal safety by maintaining fail-safe trajectories (red lines) at all times. These fail-safe trajectories are collision-free against the set of all legal behaviours (blue areas) of other traffic participants and safeguard the autonomous vehicle along its intended trajectory to safe states (grey areas).

for example, involving barrier certificates[19], Lyapunov functions[20] or automatic controller synthesis[21]. These approaches ensure that the vehicle avoids unsafe states or is kept within an invariant set of safe states[22,23] at all times. Closely related recent approaches incorporate reachability analysis to compute the set of states that a system is able to reach over time. Thus, it can be verified that unsafe states are not reached during operation[9,24–26]. However, these existing approaches are often computationally intractable, do not generalize to arbitrary traffic scenarios or do not provide the required prediction of unsafe sets in dynamic environments.

In the context of autonomous driving, the time-variant unsafe sets are commonly defined as the future occupied positions of other traffic participants, which can be obtained by motion prediction[27]. Existing prediction approaches usually compute a countable set of most likely behaviours by applying probabilistic[28–30] or machine learning methods[31–33]. The safety of autonomous vehicles is guaranteed only if no traffic participant deviates from the few predicted behaviours, but such deviations often occur in real traffic. By incorporating reachability analysis, predictions are able to consider an infinite number of possible future behaviours of dynamic obstacles[9,34–37]. Yet, allowing for all dynamically feasible behaviours of other traffic participants overly limits the manoeuvrability of the autonomous vehicle. Therefore, our reachability-based prediction only considers behaviours that are dynamically feasible in the road network and that do not violate a set of formalized traffic rules (blue areas, Fig. 1).

The motion planner for fail-safe trajectories must cope with small and convoluted solution spaces. Commonly used trajectory planning techniques either discretize the input or state space of the autonomous vehicle[38,39] or apply variational techniques in continuous space[40–42]. The former methods suffer from discretization effects, such that narrow passageways in the solution space may not be found[43] or safe terminal states may not be reached[44]. Although variational-based methods overcome these limitations, the non-convexity of the motion planning problem due to nonlinear vehicle dynamics and collision avoidance poses a major challenge. As a result, variational-based techniques are often computationally complex[45–47] or must be guided through the solution space to work in dense traffic situations[48,49], for example, by specifying driving corridors that represent temporal tactical decisions, such as overtaking an obstacle on the left or right. Approaches for obtaining driving corridors generally do not consider the dynamics of the autonomous vehicle[50–52] and thus may not be able to reason about the drivability of driving corridors. Our approach combines reachability analysis with convex optimization to determine drivable fail-safe trajectories within dynamics-aware driving corridors in arbitrary traffic scenarios (fail-safe trajectories, Fig. 1).

Results

Our verification technique ensures legal safety over consecutive verification cycles. A new verification cycle $c \in \mathbb{N}_+$ begins whenever an intended trajectory I_c is provided by the intended trajectory planner of the existing motion planning framework, where c is incremented by one for each received intended trajectory. The autonomous vehicle can only start executing a new intended trajectory I_c that is starting at t_c if I_c is successfully verified as legally safe. A trajectory is legally safe if it (1) is collision-free against the predicted occupancy sets (that is, occupied positions) that result from all legal behaviours of other traffic participants and (2) leads the autonomous vehicle to a safe terminal state.

Typically, the time horizon T_{I_c} of I_c is several seconds for planning anticipatory motions. However, the predicted occupancy sets of the surrounding traffic participants become increasingly large for longer time horizons due to growing uncertainties regarding their future behaviours. Thus, I_c is often not safe over its entire time horizon T_{I_c}. For the safety verification (Fig. 2a), we therefore do not consider the entire intended trajectory I_c, but only a short part of I_c lasting from t_c until $t_c + \Delta_c^{\text{safe}}$, where $\Delta_c^{\text{safe}} \in \mathbb{R}_+$. We regard this part of I_c as legally safe and refer to it as I_c^{safe} if it is collision-free against the predicted occupancy sets within its entire time duration Δ_c^{safe}. Because I_c^{safe} does not ensure that the autonomous vehicle remains legally safe for $t > t_c + \Delta_c^{\text{safe}}$, we compute a consecutive fail-safe trajectory F_c (the index of F indicates the corresponding intended trajectory I). The fail-safe trajectory F_c needs to smoothly continue I_c^{safe}, be collision-free against the predicted occupancy sets for its entire time horizon T_{F_c}, and transition the autonomous vehicle to a standstill in safe areas. We say that I_c is verified successfully if I_c^{safe} and F_c exist and are computed prior to t_c. The concatenation of I_c^{safe} and F_c represents the verified trajectory and is denoted as $I_c^{\text{safe}} \parallel F_c$.

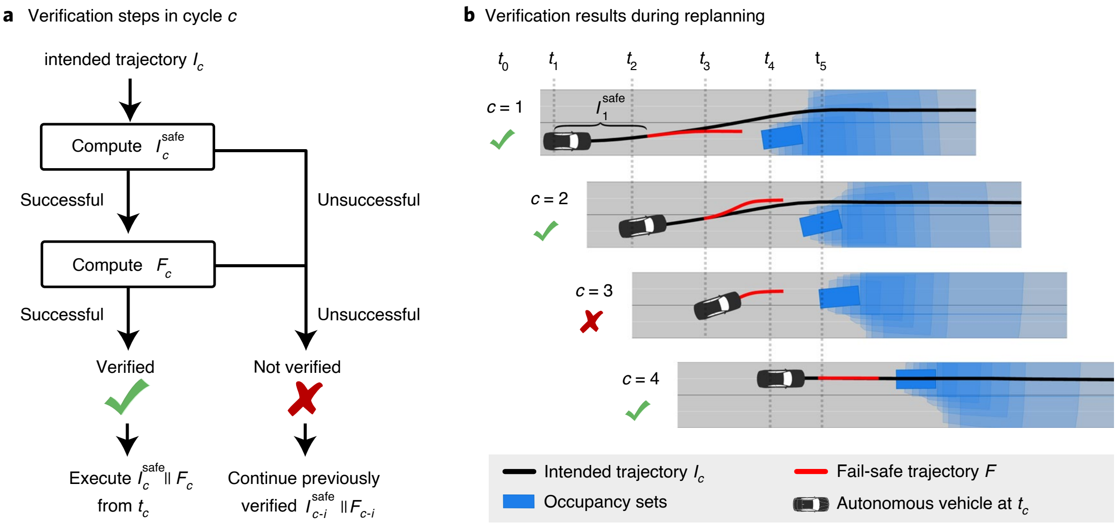
a Verification steps in cycle c
intended trajectory I_c
Compute I_c^{safe}
Successful
Unsuccessful
Compute F_c
Successful
Unsuccessful
Verified
Not verified
Execute $I_c^{\text{safe}} \parallel F_c$ from t_c
Continue previously verified $I_{c-i}^{\text{safe}} \parallel F_{c-i}$
b Verification results during replanning
t_0
t_1
t_2
t_3
t_4
t_5
$c = 1$
I_1^{safe}
$c = 2$
$c = 3$
$c = 4$
Intended trajectory I_c
Fail-safe trajectory F
Occupancy sets
Autonomous vehicle at t_c

Fig. 2 | Verification during replanning. a, In each verification cycle c, the given intended trajectory I_c is verified by computing the safe part I_c^{safe} and the fail-safe trajectory F_c. **b**, If the verification result of cycle c is successful (as in $c \in \{1, 2, 4\}$), the verified trajectory $I_c^{\text{safe}} \parallel F_c$ is executed starting at t_c. If the verification result is unsuccessful (as in $c = 3$), the verified trajectory I_{c-i}^{safe} and F_{c-i} of a previous successful verification cycle $c - i$ is executed until a new intended trajectory is successfully verified again (as in $c = 4$).

Let us explain the verification procedure during replanning using Fig. 2. Initially, at t_0, we assume that the autonomous vehicle is in a safe state (for example, parked). Immediately after the autonomous vehicle successfully verifies a given intended trajectory I_1 in verification cycle $c = 1$ (that is, I_1^{safe} and F_1 are obtained), the vehicle is allowed to engage in the autonomous driving mode at time t_1 and starts executing I_1^{safe} of the verified trajectory $I_1^{\text{safe}} \parallel F_1$ (see the result of $c = 1$ in Fig. 2b). The intended trajectory planner can then provide new intended trajectories I_c, $c > 1$, for verification. If a new trajectory I_c is successfully verified, the autonomous vehicle can transition from the previously verified trajectory to I_c^{safe} of the new verified trajectory $I_c^{\text{safe}} \parallel F_c$ at time t_c (see Fig. 2a and the result of $c \in \{2, 4\}$ in Fig. 2b). If the intended trajectory I_c cannot be verified, the most recently verified trajectory $I_{c-i}^{\text{safe}} \parallel F_{c-i}$ of cycle $c - i$, $i \in \{1, \ldots, c - 1\}$, continues to be executed (see Fig. 2a and the result of $c = 3$ in Fig. 2b). While moving along $I_{c-i}^{\text{safe}} \parallel F_{c-i}$, the fail-safe trajectory F_{c-i} is only executed if no new intended trajectory can be successfully verified before the final time of I_{c-i}^{safe}. This previously verified trajectory $I_{c-i}^{\text{safe}} \parallel F_{c-i}$ remains collision-free as long as other traffic participants do not violate traffic rules, because our set-based prediction has already anticipated all their legal future behaviours. Thus, legal safety is ensured regardless of the verification result.

Experiments on real data. For all verification cycles c in our experiments, the starting time of fail-safe trajectories F_c is equal to the starting time of the next intended trajectory I_{c+1}, that is, $t_c + \Delta_c^{\text{safe}} = t_{c+1}$ (see result for $c = 2$ in Fig. 2b). This is achieved by choosing a constant replanning rate $\Delta t = t_{c+1} - t_c$ (meaning that new intended trajectories should be executed at rate Δt) that is set to the constant duration of I_c^{safe} as $\Delta t = \Delta_c^{\text{safe}}$ for all c. Consequently, when executing a verified trajectory $I_c^{\text{safe}} \parallel F_c$, the transition to the fail-safe trajectory F_c may only occur at t_{c+1}. Thus, in each time interval $[t_c, t_{c+1}]$, the autonomous vehicle either executes I_c^{safe} completely or a part of F_{c-i} of a previously verified $I_{c-i}^{\text{safe}} \parallel F_{c-i}$. In other words, only if the current verification result is not successful do the autonomous vehicles transition from the safe part of an intended trajectory to a fail-safe trajectory.

In urban environments, most accidents occur at intersections and with pedestrians[53]. To demonstrate that our proposed verification technique allows autonomous vehicles to handle these crucial cases, we created two scenarios by recording real traffic with a BMW 7 series vehicle. By post-processing the real-world recordings, as described in the Supplementary Information, and applying our verification technique offline, we obtained the results presented below. For each of the two scenarios we illustrate an overview of the traffic situation using recorded images from the BMW 7 series vehicle and show the verification results of selected verification cycles c (Figs. 3 and 4). In addition, we demonstrate for both scenarios that our method guarantees legal safety for arbitrary intended trajectory planners (Fig. 5). In the Supplementary Information, we further provide a scenario illustrating safe lane changes (where the third most accidents occur[53]), further results including videos, detailed computation times (177 ms on average), all used parameters and software to visualize the verification results for all verification cycles.

Scenario I: left-turn at an urban intersection. In countries where vehicles drive on the right (we apply this throughout this Article), left turns at intersections are among the most hazardous manoeuvres, because the autonomous vehicle must consider the right of way of oncoming vehicles and yield to potential cyclists in their dedicated lane (Fig. 3a). The behaviour of oncoming vehicles or cyclists may change rapidly over time. For example, vehicles may accelerate or decelerate, and cyclists may even stop and dismount, which increases the uncertainty about the future evolution of the traffic scenario. Under all circumstances, the autonomous vehicle must yield to oncoming traffic while not disrupting the traffic flow due to overly conservative behaviour.

Our method accomplishes this challenge by safeguarding the opportunistic intended trajectory with fail-safe trajectories that (1) comply with the right of way and (2) never stop the autonomous vehicle in the intersection area. Because our prediction accounts for all legal behaviours of other traffic participants, our verification technique can decide whether a left turn manoeuvre can be completed before oncoming traffic can enter the intersection. Thus, the autonomous vehicle automatically respects the right of way.

As illustrated in Fig. 3b at $t_1 = 0$ s, the autonomous vehicle first approaches the intersection along its intended trajectory, that is, I_c^{safe}, $c \in \{1, \ldots, 4\}$, is executed. From $t_5 = 2.4$ s to $t_{10} = 5.4$ s, our

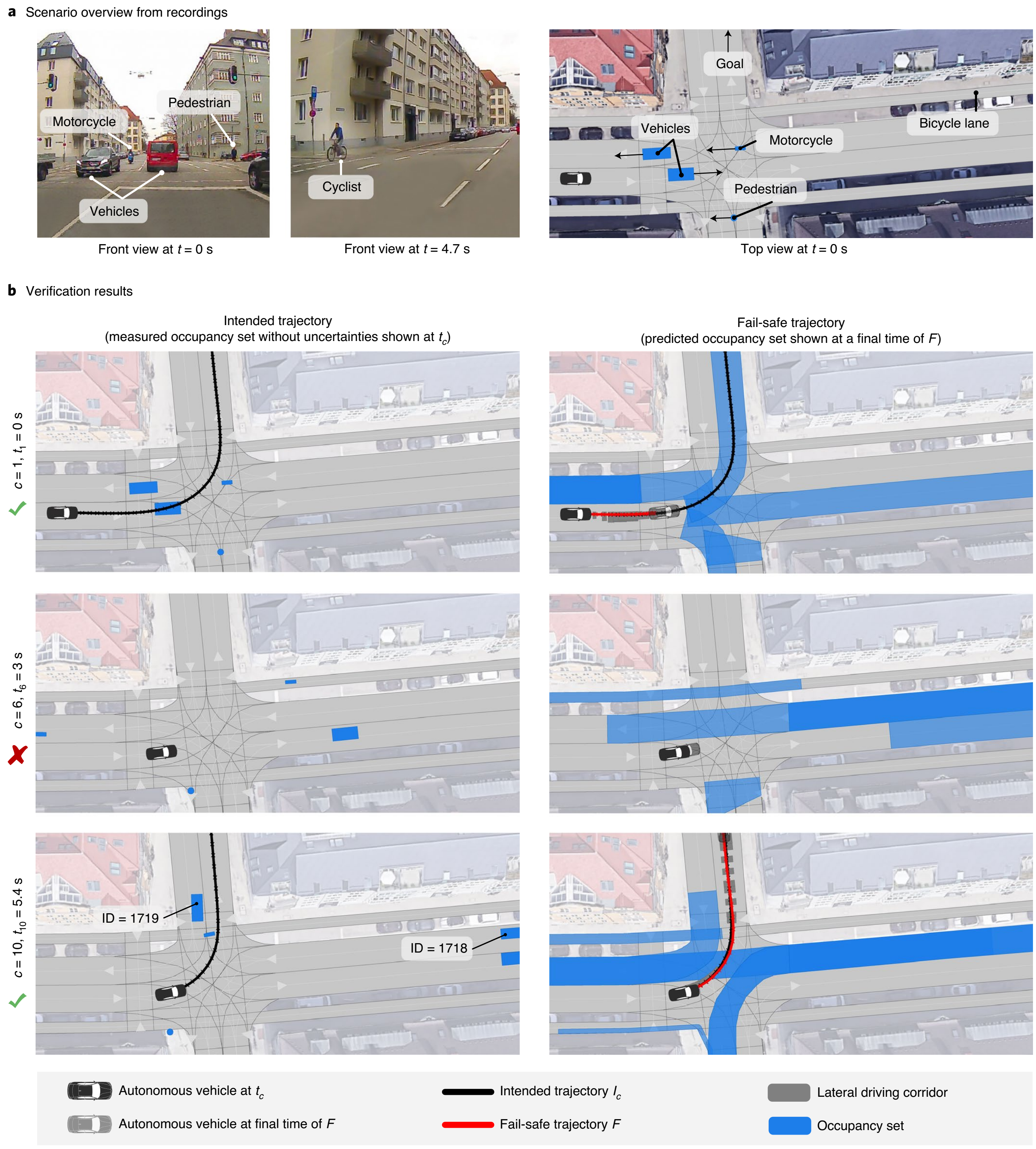

Fig. 3 | Results of Scenario I (urban intersection). a, Camera images and top view of the scenario. **b**, Verification results of selected verification cycles c. The intended trajectory I_c is only shown if it is successfully verified. Credit: Google, GeoBasis-DE/BKG (satellite images).

approach automatically detects that the intended trajectories lead to an unsafe situation in which a collision with the oncoming vehicle within the intersection area cannot be excluded before the cyclist has definitely passed. The fail-safe trajectory thus stops the autonomous vehicle at the intersection (see fail-safe trajectory at $t_6 = 3$ s in Fig. 3b). Immediately after the cyclist has passed, our verification technique successfully verifies an intended trajectory and the autonomous vehicle continues its left turn before oncoming traffic, as shown in Fig. 3b at $t_{10} = 5.4$ s. Note that, in this figure, the fail-safe trajectory overlays the occupancy sets, because the occupancy sets are shown at the final time of the fail-safe trajectory (see Supplementary Fig. 8 for the occupancy sets at intermediate times). Figure 3b also demonstrates that our prediction incorporates traffic rules. Consider the occupancy set of the oncoming vehicle with

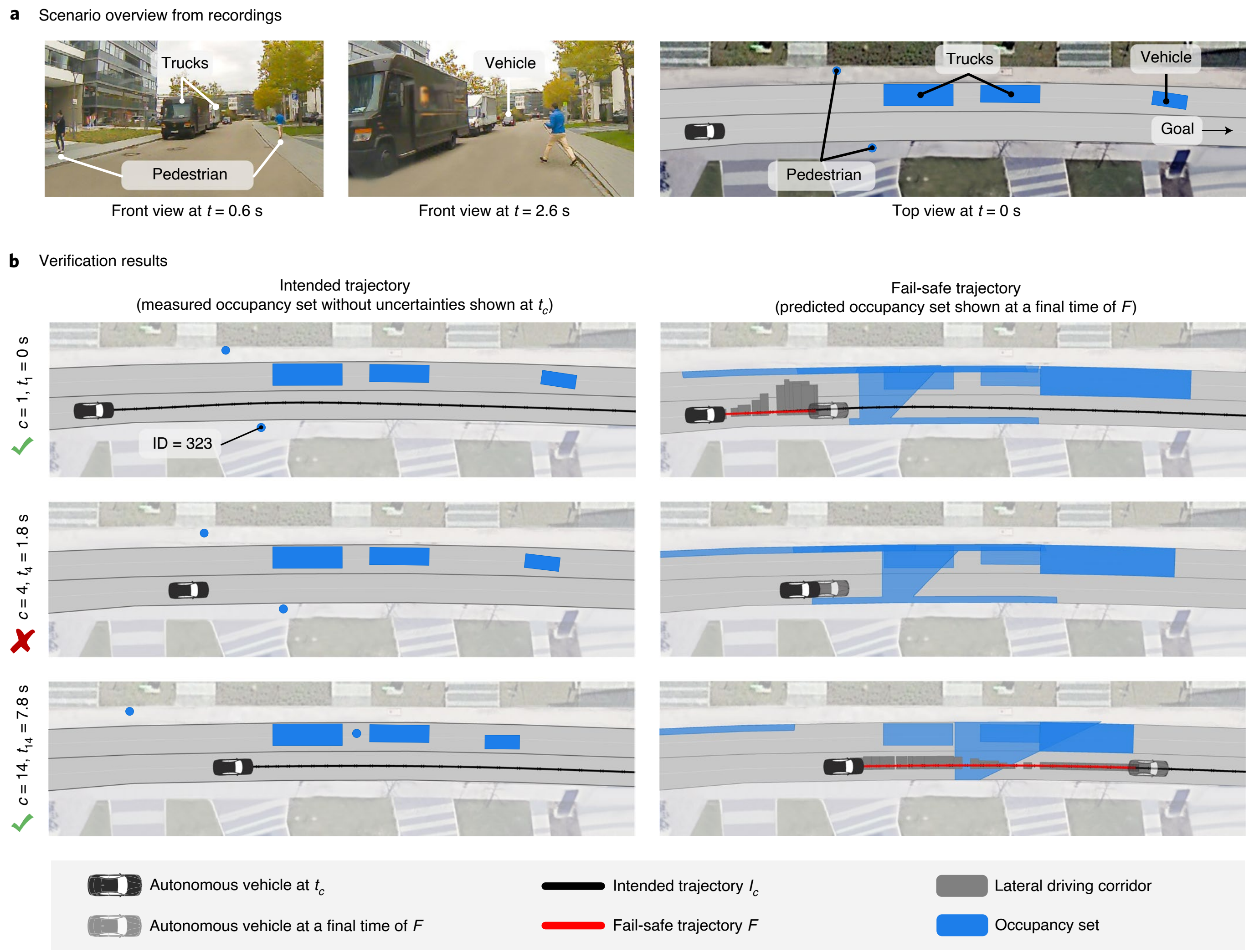
a Scenario overview from recordings
Trucks
Pedestrian
Front view at $t = 0.6$ s
Vehicle
Front view at $t = 2.6$ s
Trucks
Vehicle
Goal
Pedestrian
Top view at $t = 0$ s
b Verification results
Intended trajectory (measured occupancy set without uncertainties shown at t_c)
Fail-safe trajectory (predicted occupancy set shown at a final time of F)
$c = 1$, $t_1 = 0$ s
ID = 323
$c = 4$, $t_4 = 1.8$ s
$c = 14$, $t_{14} = 7.8$ s
Autonomous vehicle at t_c
Autonomous vehicle at a final time of F
Intended trajectory I_c
Fail-safe trajectory F
Lateral driving corridor
Occupancy set

Fig. 4 | Results of Scenario II (jaywalking pedestrian). **a**, Camera images and top view of the scenario. **b**, Verification results of selected verification cycles c. The intended trajectory I_c is only shown if it is successfully verified. Credit: Google, GeoBasis-DE/BKG (satellite images).

ID 1718 at $t_{10} = 5.4$ s. The legal safe distance forbids vehicles to turn after the autonomous vehicle in a way that obstructs the autonomous vehicle. Therefore, the vehicle with ID 1718 is only allowed to continue straight or turn left, but may not yet turn right.

Scenario II: jaywalking pedestrian. Vulnerable road users pose a special challenge to autonomous vehicles, because they often exhibit unexpected changes in behaviour. In particular, pedestrians can quickly change their walking direction, which makes it difficult for autonomous vehicles to react in time. Even though it is illegal for pedestrians to jaywalk, that is, to cross the road in the presence of traffic, pedestrians are occasionally inattentive and cross directly in front of passing vehicles. If the prediction of the autonomous vehicle does not include this behaviour, a fatal accident could occur.

In the first verification cycle $c = 1$ presented in Fig. 4, the pedestrian with ID 323 (in a blue jacket) is walking on the sidewalk and is only looking at his cell phone (Fig. 4a). To anticipate that this inattentive pedestrian may jaywalk, we broaden the set of considered legal behaviours for this pedestrian by relaxing the constraints in its prediction. As a result, the autonomous vehicle computes the future occupancies of this pedestrian for both crossing the road and walking partially on the road parallel to the sidewalk (see occupancy set in Fig. 4b for the fail-safe trajectory at $t_1 = 0$ s; note that occupancy sets of pedestrians are not visualized outside of the road). The resulting fail-safe trajectory F_1 (starting at t_2) ensures that the autonomous vehicle remains behind the pedestrian.

In the next verification cycles $c \in \{2, 3, 4\}$, the autonomous vehicle cannot verify the new intended trajectories. In fact, each intended trajectory collides with the jaywalking pedestrian. Thus, by automatically executing the first computed fail-safe trajectory F_1, the autonomous vehicle slows down to avoid a collision with the pedestrian with ID 323 (see $t_4 = 1.8$ s in Fig. 4b). After the pedestrian crosses, the autonomous vehicle accelerates to the desired velocity, and the fail-safe trajectory implies that the autonomous vehicle is able to pass before the pedestrian may walk back towards the lane of the autonomous vehicle (see $t_{14} = 7.8$ s in Fig. 4b).

As demonstrated in this scenario, our verification technique offers its users, such as mobility providers, the flexibility to define the legal behaviours differently for specific types of traffic participants. For example, when driving past a school, one may wish to anticipate that any child or even any pedestrian may cross the road.

Legal safety for arbitrary intended trajectories. We apply our verification technique to three different intended trajectory planners (for details see Supplementary Information):

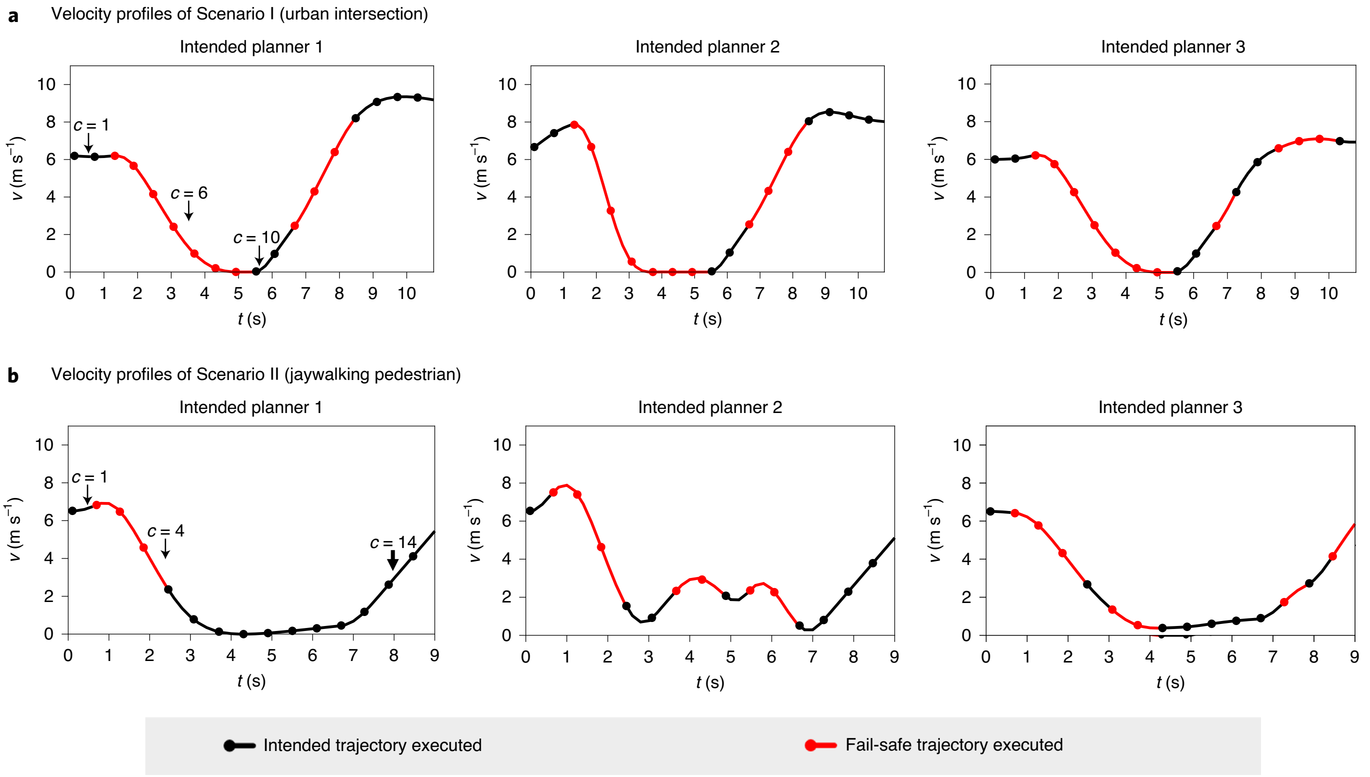

Fig. 5 | Results of the verification technique with different intended planners. a, Executed velocity profiles in Scenario I (results of cycles $c \in \{1, 6, 10\}$ are labelled). **b**, Executed velocity profiles in Scenario II (results of cycles $c \in \{1, 4, 14\}$ are labelled).

- Planner 1 uses continuous optimization to plan trajectories that are collision-free with regard to the most likely behaviour of other traffic participants. This planner is also used as an intended trajectory planner for the previous results of Scenarios I and II.
- Planner 2 is based on Planner 1 with the modification that other traffic participants are ignored. With this planner, we mimic a reinforcement learning approach that has not yet learned collision avoidance.
- Planner 3[39] samples in a discrete state space to plan trajectories that are collision-free with regard to the most likely behaviour of other traffic participants.

Figure 5 illustrates the velocity profiles of the autonomous vehicle in Scenarios I and II for each intended trajectory planner. In Scenario I, our verification technique intervenes independently of the applied intended trajectory planner so that the autonomous vehicle stops in front of the intersection (Fig. 5a). Although Planner 2 is not aware of other traffic participants, our verification technique enables the autonomous vehicle to safely turn left. Because Planner 2 tries to reach the desired velocity (8 m s^{-1}) more aggressively than Planners 1 and 3 (see the results of verification cycles $c \in \{1, 2\}$ in Fig. 5a), the subsequently executed fail-safe trajectories cause a rapid deceleration of the autonomous vehicle (peak, −6 m s^{-2}) (see the results of verification cycles $c \in \{3, \ldots, 8\}$ for Intended Planner 2 in Fig. 5a). However, the execution of fail-safe trajectories for Planner 2 causes only a short delay, as the stopping time at the intersection is less than 2 s.

In Scenario II, the intended trajectory planners are not aware of the pedestrian's intention to jaywalk. Therefore, fail-safe trajectories are executed to slow down the autonomous vehicle (see the results of verification cycles $c \in \{2, 3, 4\}$ in Fig. 5b) until Planners 1 and 3 react to the pedestrian. Planner 2 requires permanent guidance to avoid a collision with the pedestrian. Although the type of executed trajectory, that is, I_c^{safe} or F_{c-i}, continuously alternates, the average velocity of the autonomous vehicle with Planner 2 is 5% higher than that with Planner 1 (6.36 m s^{-1} and 6.09 m s^{-1}, respectively).

In summary, we are able to guarantee legal safety for different intended trajectory planners, even when using a planner that ignores other traffic participants. Furthermore, the resulting velocity profiles are smooth and continuous, as fail-safe trajectories are planned with full consideration of the vehicle's dynamics.

Discussion

Certification is the main obstacle to achieving commercial success with the proposed verification technique. Regulatory guidelines have already been prepared for various domains, such as railway systems, industrial robots and aviation systems, but only limited regulations exist for motion planning of autonomous vehicles (for example, ISO 26262 and ISO 21448). We have prepared the ground for certification by formulating legal safety and presenting a verification technique that ensures that this specification is met during operation of the autonomous vehicle. Moreover, the safety guarantees are maintained when adapting our considered set of traffic rules to new requirements. If legal safety becomes a recognized standard for autonomous vehicles, mobility providers can certify our proposed verification technique for usage in their vehicles. As a result, we expect that societal trust in autonomous vehicles will increase and that testing efforts can be significantly reduced, even if motion planning frameworks for generating intended trajectories are changed.

Legal safety is a promising novel safety approach inspired by traffic regulations that is suitable for certification. Related concepts, such as responsibility-sensitive safety[54], not-at-fault driving[26] and compositional and contract-based verification[55], share our premise to avoid (self-inflicted) accidents, but differ substantially to our proposed solution. Responsibility-sensitive safety assumes that other traffic participants act according to common-sense rules

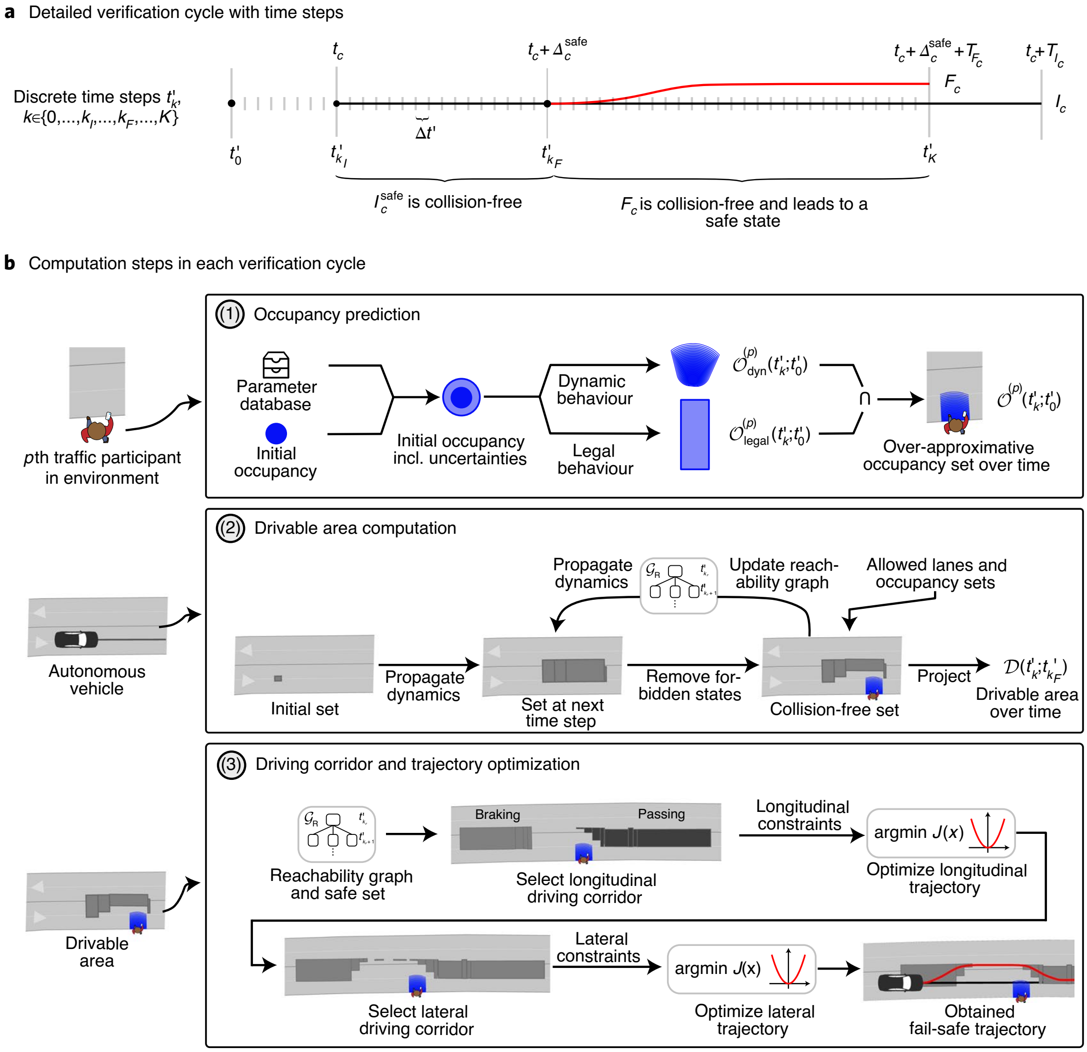
a Detailed verification cycle with time steps
Discrete time steps t'_k, $k\in\{0,\ldots,k_I,\ldots,k_F,\ldots,K\}$
t_c
$t_c+\Delta_c^{\mathrm{safe}}$
$t_c+\Delta_c^{\mathrm{safe}}+T_{F_c}$
$t_c+T_{I_c}$
F_c
I_c
$\Delta t'$
t'_0
t'_{k_I}
t'_{k_F}
t'_K
I_c^{safe} is collision-free
F_c is collision-free and leads to a safe state
b Computation steps in each verification cycle
pth traffic participant in environment
(1) Occupancy prediction
Parameter database
Initial occupancy
Initial occupancy incl. uncertainties
Dynamic behaviour
Legal behaviour
$\mathcal{O}^{(p)}_{\mathrm{dyn}}(t'_k;t'_0)$
$\mathcal{O}^{(p)}_{\mathrm{legal}}(t'_k;t'_0)$
∩
$\mathcal{O}^{(p)}(t'_k;t'_0)$
Over-approximative occupancy set over time
Autonomous vehicle
(2) Drivable area computation
Propagate dynamics
$\mathcal{G}_R$
Update reachability graph
Allowed lanes and occupancy sets
Initial set
Propagate dynamics
Set at next time step
Remove forbidden states
Collision-free set
Project
$\mathcal{D}(t'_k;t'_{k_F})$
Drivable area over time
Drivable area
(3) Driving corridor and trajectory optimization
Reachability graph and safe set
Braking
Passing
Select longitudinal driving corridor
Longitudinal constraints
argmin $J(x)$
Optimize longitudinal trajectory
Select lateral driving corridor
Lateral constraints
argmin $J(x)$
Optimize lateral trajectory
Obtained fail-safe trajectory

Fig. 6 | Computation steps of the verification technique. a, Time discretization t'_k in one verification cycle c. **b**, Overview of the computation steps for verifying an arbitrary intended trajectory. (1) We compute occupancy sets, that is, all legally occupied positions of other traffic participants over time. (2) The drivable area of the autonomous vehicle is computed to determine fail-safe manoeuvres. (3) Longitudinal and lateral driving corridors are selected from the reachability graph, and longitudinal and lateral trajectories are optimized such that a fail-safe trajectory is obtained.

and defines appropriate responses by the autonomous vehicle based on safe distances. However, despite the execution of appropriate responses, self-inflicted accidents cannot be excluded, because other traffic participants may behave differently than expected. Our approach addresses this problem by considering all legal behaviours. Not-at-fault driving computes a single trajectory that is split into moving, braking and stopped phases and is provably collision-free against a given prediction. By contrast, we allow intended trajectories to be planned independently of fail-safe trajectories, for example, using a most likely prediction to optimize comfort. In ref. [55], a finite number of offline-verified, local models are fitted online to the current traffic situation. However, this approach may result in unsafe behaviours if no valid composition of these local models can be found for the current situation. Our verification technique evaluates the safety of situations online and always provides fail-safe trajectories to eliminate self-inflicted accidents. The detailed computation steps of our verification technique are described in the Methods and are visualized in Fig. 6.

Methods

Formal verification is often believed to cause performance drops (for example, lower average velocities resulting in longer travel times) and conservative behaviour in robotic systems[56,57]. However, we believe that autonomous vehicles can offer high performance and ensure legal safety at the same time. This has motivated us to improve on our previous work on set-based predictions[58–60], fail-safe trajectory planning[61] and trajectory planning using reachable sets[62]. Further to our previous work, we present the following innovations:

1. Our proposed verification technique ensures legal safety in complex traffic scenarios and in a computationally efficient way. In particular, by embedding driving corridors[62] into fail-safe trajectory planning[61], we generalize the computation of possible fail-safe manoeuvre options to different traffic situations and can consider multiple safe terminal sets.
2. On various urban scenarios that have been recorded in real traffic including measurement uncertainties, the applicability of the proposed verification technique is demonstrated. In addition, our results indicate that non-conservative driving behaviour can be achieved despite the over-approximative, set-based prediction.
3. The temporal interplay over subsequent verification cycles of our verification technique with the intended trajectory planner of the autonomous vehicle is presented in detail.
4. Further experiments with three different intended trajectory planners validate that our verification technique is able to ensure legal safety for arbitrary intended trajectory planners.

In the following paragraphs, we present the inputs of our verification technique, preliminaries for the reachability analysis, an overview of the algorithmic steps and the safety guarantees for our verification technique. Additional details are provided in the Supplementary Information.

Inputs of the verification technique. Our verification technique is integrated between the motion planning layer and the control layer of the autonomous vehicle (see planning frameworks in refs. [63,64]). In each verification cycle c, our verification technique receives as inputs the intended trajectory I_c and the environment model. The intended trajectories must be kinematically feasible and branch off the previously verified trajectory $I_{c-i}^{\text{safe}} \parallel F_{c-i}$. The environment model must contain the lanes of the road, pedestrian crossings and areas in which the autonomous vehicle is not allowed to stop, which are used to obtain the designated safe areas. For all safety-relevant traffic participants, the environment model must contain their type (that is, vehicle, motorcycle, bicycle or pedestrian) and their current states (that is, a set containing the exact state and bounded measurement uncertainties). If the type of traffic participant is unknown or uncertain, our verification technique can predict the set of future behaviours for all possible types in parallel.

Preliminaries of the verification technique. The motion of the pth traffic participant is governed by the differential equation $\dot{x}^{(p)}(t) = f^{(p)}\left(x^{(p)}(t), u^{(p)}(t)\right)$, where $x^{(p)}$ is the state and $u^{(p)}$ is the input. The admissible states and inputs are bounded by the respective sets $\mathcal{X}^{(p)}(t) \subset \mathbb{R}^{n^{(p)}}$ and $\mathcal{U}^{(p)}(t) \subset \mathbb{R}^{m^{(p)}}$. A possible solution of the differential equation at time t is denoted by $\chi^{(p)}\left(t; x^{(p)}(\tau_0), u^{(p)}(\cdot)\right)$, when starting at state $x^{(p)}(\tau_0) \in \mathcal{X}_0^{(p)}$, where $\mathcal{X}_0^{(p)}$ is the set of states at an initial time τ_0 including measurement uncertainties, and using input trajectory $u^{(p)}(\cdot)$. The reachable set $\mathcal{R}^{\mathrm{e}^{(p)}}(t; \tau_0) \subseteq \mathcal{X}^{(p)}(t)$ describes the set of states that are reachable by the pth traffic participant at a certain point in time $t \geq \tau_0$ when starting in $\mathcal{X}_0^{(p)}$ and applying all admissible inputs $\mathcal{U}^{(p)}(t)$:

$$\begin{aligned}\mathcal{R}^{\mathrm{e}^{(p)}}(t; \tau_0) = \Big\{\chi^{(p)}\left(t; x^{(p)}(\tau_0), u^{(p)}(\cdot)\right) \,\Big|\, & x^{(p)}(\tau_0) \in \mathcal{X}_0^{(p)}, \forall \tilde{\tau} \in [\tau_0, t] : \\ & \chi^{(p)}\left(\tilde{\tau}; x^{(p)}(\tau_0), u^{(p)}(\cdot)\right) \in \mathcal{X}^{(p)}(\tilde{\tau}), u^{(p)}(\tilde{\tau}) \in \mathcal{U}^{(p)}(\tilde{\tau})\Big\}\end{aligned} \tag{1}$$

For brevity, we omit the superscript (p) when referring to the autonomous vehicle. In each verification cycle c, we compute the reachable set of other traffic participants to predict their future movement and that of the autonomous vehicle to obtain its drivable area.

As illustrated in Fig. 6a, we introduce the discrete points in time t'_k for each verification cycle c, where $k \in \{0, \ldots, k_I, \ldots, k_F, \ldots K\} \subseteq \mathbb{N}_0$; for brevity, the notation of t'_k does not reflect its dependency on c. Time t'_0 is the initial time of the prediction, that is, the point in time at which the most recently available environment model has been recorded. Time t'_{k_I} corresponds to the start time of the intended trajectory I_c (that is, $t'_{k_I} = t_c$), t'_{k_F} corresponds to the start time of the fail-safe trajectory F_c (that is, $t'_{k_F} = t_c + \Delta_c^{\text{safe}}$) and t'_K corresponds to the final time of the fail-safe trajectory (that is, $t'_K = t_c + \Delta_c^{\text{safe}} + T_{F_c}$). Without loss of generality, we assume that the times t'_k are multiples of the time step size $\Delta t' \in \mathbb{R}_+$, that is, $t'_k = t'_0 + k\Delta t'$.

Recall that we set Δ_c^{safe} to the replanning rate Δt in our experiments. To minimize the interventions of our verification technique, that is, how often a fail-safe trajectory is executed, the duration Δ_c^{safe} can be dynamically adjusted to optimize the length of I_c^{safe} as described in ref. [65]. To avoid that new intended trajectories cannot be verified solely due to a timeout, intended trajectories I_c should be provided prior to $t_c - \Delta^{\text{verify}}$, where $\Delta^{\text{verify}} \in \mathbb{R}_+$ is the required computation time of our verification method.

Occupancy prediction. The goal in the first step of our verification technique is to over-approximate the area $\mathcal{L}^e(t)$ that exactly encloses the occupied positions of the surrounding traffic participants for all their legal behaviours. Therefore, we first compute all dynamically feasible behaviours and subsequently remove illegal behaviours.

All dynamically feasible behaviours of other traffic participants are obtained using reachability analysis as defined in equation (1). For each pth traffic participant, the environmental model provides the initial states $\mathcal{X}_0^{(p)}$ at t'_0, which are described by a set due to measurement uncertainties (Fig. 6b, step (1)). The dynamics of each traffic participant are abstracted by a second-order integrator model with bounded velocities and accelerations. We compute the reachable set $\mathcal{R}^{(p)}(t; t'_0)$ as a tight over-approximation of the exact reachable set, that is, $\mathcal{R}^{(p)}(t; t'_0) \supseteq \mathcal{R}^{e^{(p)}}(t; t'_0)$, and only for the position domain to allow for an efficient computation. For collision checks with planned trajectories of the autonomous vehicle, we introduce $\mathcal{O}_{\text{dyn}}^{(p)}(t; t'_0)$ as the dynamics-based occupancy set resulting from the over-approximative reachable set $\mathcal{R}^{(p)}(t; t'_0)$ by considering the dimensions of the pth traffic participant (Fig. 6b, step (1)).

Next, we remove behaviours that are not allowed according to traffic rules. Therefore, we formalize a set of traffic rules that is most relevant for motion planning (and which can be easily extended). Let $v^{(p)}$ and $a^{(p)}$ denote the velocity and acceleration of the pth predicted traffic participant, respectively, and $\diamondsuit^{\text{veh}}$ denotes that the parameter $\diamondsuit \in \{\overline{v}, \overline{a}\}$ bounding the velocity or acceleration is applicable for vehicles and motorcycles, while $\diamondsuit^{\text{cyc}}$ is for bicycles and $\diamondsuit^{\text{ped}}$ is for pedestrians (the values of the parameters are stored in a database generated offline, can be updated online, and are provided in the Supplementary Information). The considered traffic rules for vehicles, motorcycles and bicycles are as follows:

- Maximum velocity is bounded (article 13.2 of ref. [7]): $v^{(p)} \leq v_{\text{limit}} f_S^{(p)}$, where v_{limit} is the legal speed limit of the road and $f_S^{(p)} \geq 1$ is a parameterized speeding factor to consider slight over-speeding. If no speed limit is available, such as for bicycles, $v^{(p)} \leq \overline{v}^{\text{veh/cyc}}$.
- Driving backward is not allowed (article 14.2 of ref. [7]): $v^{(p)} \geq 0$.
- Absolute acceleration is bounded (due to tyre friction): $|a^{(p)}| \leq \overline{a}^{\text{veh/cyc}}$.
- Leaving the road is forbidden (article 14.1 of ref. [7]).
- A safe distance to the autonomous vehicle must be maintained when driving behind it or merging in front of it (articles 13.5 and 11.2d of ref. [7]).
- Changing lanes is only allowed if the new lane has the same driving direction as the previous one (article 11.2c of ref. [7]).

Note that, according to article 11.2c of ref. [7], overtaking in a lane not appropriate to the direction of traffic is only allowed if not endangering or interfering with oncoming traffic. Because such a legal overtaking manoeuvre does not interfere with the motion planning of the autonomous vehicle, we neglect it in our prediction without compromising legal safety.

Although pedestrians are generally not allowed to obstruct vehicular traffic, for example, to jaywalk (article 7.1 of ref. [7]), vehicles are required to take precautions to avoid endangering pedestrians (article 21.1 of ref. [7]). Thus, the considered traffic rules for pedestrians are as follows:

- Absolute velocity is bounded (for example, based on ISO 13855): $|v^{(p)}| \leq \overline{v}^{\text{ped}}$.
- Absolute acceleration is bounded (due to physical capabilities): $|a^{(p)}| \leq \overline{a}^{\text{ped}}$.
- Entering the road is forbidden (articles 7.1 and 20.2 of ref. [7]) except
  - on pedestrian crossings (articles 20.6b and 21.2 of ref. [7])
  - when walking toward the road; then, crossing the road is allowed perpendicularly with a deviation of angle α based on the current heading of the pedestrian (articles 20.6c,d of ref. [7])
  - when walking parallel to the road; then, occupying the strip of the road edge with a width of d_{slack} is allowed, for example, to avoid obstacles on the sidewalk (articles 20.2a, 20.3 and 20.4 of ref. [7]).

In summary, our set of traffic rules either constrains the dynamics of other traffic participants (for example, their maximum velocity), which are considered by $\mathcal{O}_{\text{dyn}}^{(p)}(t; t'_0)$, or constrains the allowed regions in the environment (for example, certain lanes or pedestrian crossings), which are given by the environment model and are denoted by $\mathcal{O}_{\text{legal}}^{(p)}(t; t'_0)$. The resulting over-approximative occupancy set of the pth traffic participant is $\mathcal{O}^{(p)}(t; t'_0) = \mathcal{O}_{\text{dyn}}^{(p)}(t; t'_0) \cap \mathcal{O}_{\text{legal}}^{(p)}(t; t'_0)$ (Fig. 6b, step (1)). To verify that I_c^{safe} and F_c are collision-free, we compute the occupancy sets for consecutive time intervals $[t'_k, t'_{k+1}]$ until the final time of F_c, that is, $\forall\, k \in \{k_I, \ldots, K\}$. Note that the time intervals $[t'_k, t'_{k+1}]$ can be of different duration for each k, for example, in case I_c^{safe} and F_c are discretized differently. The predicted occupancy sets of all traffic participants are given by $\mathcal{L}([t'_k, t'_{k+1}]) = \bigcup_p \bigcup_{t \in [t'_k, t'_{k+1}]} \mathcal{O}^{(p)}(t; t'_0)$.

Note that, regardless of how many traffic rules we consider, our prediction always over-approximates the exact set of all legal behaviours, that is, $\mathcal{L}(t) \supseteq \mathcal{L}^e(t)$. The reason is that only behaviours defined as illegal are removed from the over-approximation of all dynamically feasible behaviours. The fewer traffic rules we consider, the more cautiously the autonomous vehicle behaves, because it respects more behaviours than actually allowed according to all traffic rules. However, the autonomous vehicle definitely remains collision-free when other traffic participants adhere to all traffic rules, as prescribed by legal safety. If a collision occurs nonetheless, we can verifiably argue that another traffic participant must have violated traffic rules and that the collision is not self-inflicted by the autonomous vehicle. Nevertheless, we account for humans' tendency to violate traffic rules, such as the speed limit. Therefore, we continuously monitor whether any traffic participant performs a behaviour that is not included in the set of legal behaviours. Whenever violations are detected, this behaviour is automatically added to the prediction result; for example, if another vehicle illegally changes lanes, we no longer exclude this behaviour from our prediction of this vehicle. As a result, our verification technique will attempt to find a new fail-safe trajectory in case the previous one is no longer collision-free. Furthermore, if a traffic participant appears likely to misbehave, such behaviours can be included in our prediction by disabling the corresponding constraint, as demonstrated in Scenario II.

Drivable area computation. To obtain possible sequences of high-level fail-safe manoeuvres (for example, overtaking other vehicles on their left or right), we compute the drivable area of the autonomous vehicle at discrete points in time t'_k with $k \geq k_F$ by projecting its reachable set $\mathcal{R}^e(t'_k; t'_{k_F})$ defined in equation (1) onto the position domain (Fig. 6b, step (2)). As for the prediction of other traffic participants, we abstract the dynamics of the autonomous vehicle using two second-order integrator models in the longitudinal and lateral directions with bounded velocities and accelerations in a road-aligned coordinate system[66]. For computational efficiency, the reachable set is approximated through the union of base sets $\mathcal{B}_k^{(i)}$, $i \in \mathbb{N}_0$, such that $\mathcal{R}^e(t'_k; t'_{k_F}) \approx \bigcup_i \mathcal{B}_k^{(i)}$ holds. The base sets $\mathcal{B}_k^{(i)}$ are the Cartesian products of convex polytopes describing reachable position–velocity pairs in the longitudinal and lateral directions. We use convex polytopes, because they are closed under required set operations such as Minkowski sum, linear mapping and intersection. The projection of base sets $\mathcal{B}_k^{(i)}$ onto the position domain yields axis-aligned rectangles $\mathcal{D}_k^{(i)}$ that represent the drivable area $\mathcal{D}(t'_k; t'_{k_F}) := \bigcup_i \mathcal{D}_k^{(i)}$. The projection of the reachable set onto the position domain can be computed efficiently, because we only need to determine the minimum and maximum position coordinates of the convex polytopes of the base sets $\mathcal{B}_k^{(i)}$.

The state $x(t'_{k_F})$ of the fail-safe trajectory F_c at its start time t'_{k_F} is provided by the final state of I_c^{safe}. We enclose $x(t'_{k_F})$ with a base set such that $x(t'_{k_F}) \in \mathcal{B}_{k_F}^{(0)}$ holds. The reachable set of consecutive points in time t'_{k+1}, $k \geq k_F$, is computed as illustrated in Fig. 6b (step (2)). First, we propagate each base set $\mathcal{B}_k^{(i)}$ of the previous time step forward in time considering all admissible inputs. Second, we remove states outside the set of admissible states $\mathcal{X}(t'_{k+1})$, that is, positions in which the autonomous vehicle collides with the predicted occupancy sets $\mathcal{L}([t'_k, t'_{k+1}])$ or the area $\mathcal{Q}$ outside of the road, to obtain $\mathcal{R}(t'_{k+1}; t'_{k_F}) \approx \bigcup_j \mathcal{B}_{k+1}^{(j)}$ at time t'_{k+1}. Third, we store each base set $\mathcal{B}_{k+1}^{(j)}$ in a directed graph $\mathcal{G}_{\mathcal{R}}$. In $\mathcal{G}_{\mathcal{R}}$, each set $\mathcal{B}_{k+1}^{(j)}$ is associated with exactly one node and an edge indicates that base set $\mathcal{B}_{k+1}^{(j)}$ is reachable from $\mathcal{B}_k^{(i)}$ within one time step. The procedure is repeated until the final time step t'_K is reached.

Driving corridor and trajectory optimization. We generate drivable fail-safe trajectories through continuous optimization. As convex optimization problems can be solved efficiently with global convergence, we convexify the inherently non-convex optimization problem by separating the longitudinal and lateral motion of the autonomous vehicle. However, longitudinal motion planning requires prior knowledge on the lateral motion and vice versa, as both subsystems are dynamically coupled. To overcome this issue, we obtain driving corridors from the drivable area that provide spatio-temporal position constraints for the optimization problems. We refer to the driving corridors for longitudinal and lateral optimization as the longitudinal and lateral driving corridors, respectively. To ensure legal safety for an infinite time horizon, we constrain the driving corridors to end in a safe terminal state based on the designated safe areas, for example, a standstill in the rightmost lane sufficiently far from an intersection. As illustrated in Fig. 6b (step (3)), our motion planner first optimizes the longitudinal trajectory within a longitudinal driving corridor, followed by optimizing the lateral trajectory in a suitable lateral driving corridor. Currently, we constrain fail-safe trajectories to be kinematically feasible, collision-free with respect to road boundaries and the predicted occupancy sets, respect the speed limit and end in a safe state. Further constraints can be imposed to consider additional properties, for example, rules on overtaking or stopping at the boundaries of the field of view of the vehicle.

We represent collision avoidance constraints by a minimum and maximum value on the longitudinal or lateral positions at each point in time. To obtain these limits, we exploit that a connected set in the position domain projected onto either the longitudinal or lateral direction yields an interval. Consequently, we define a longitudinal corridor and a lateral driving corridor for fail-safe motion planning as a temporal sequence of connected sets that are subsets of the drivable area $\mathcal{D}(t'_k; t'_{k_F})$ from time t'_{k_F} to the final time t'_K.

To determine longitudinal driving corridors, we perform a search on the reachability graph $\mathcal{G}_{\mathcal{R}}$ backwards in time starting from the set of safe terminal states (Fig. 6b, step (3)). There may be multiple longitudinal driving corridors, because the drivable area can be disconnected due to surrounding traffic participants. We select the longitudinal driving corridor with the greatest cumulative drivable area from t'_{k_F} to t'_K for trajectory planning (other heuristics can also be applied). For the longitudinal trajectory optimization, we use a fourth-order integrator model with jounce as input and bounded longitudinal velocity, acceleration and jerk. In addition to the collision avoidance constraints from the boundary of the longitudinal driving corridor, the autonomous vehicle must come to a standstill at the final time t'_K. To improve comfort, we choose a quadratic cost function that minimizes acceleration, jerk and jounce as well as deviations from the desired velocity.

The computation and selection of lateral driving corridors are performed similarly to the computation and selection of longitudinal driving corridors with the addition that the connected sets of the lateral driving corridor must provide a unique passing side for each obstacle. The lateral trajectories of the autonomous vehicle are optimized with respect to a linearized kinematic single-track model with limits on the steering actuators. Analogously to planning in the longitudinal direction, the position constraints for collision avoidance are obtained from the boundaries of the lateral driving corridor. We select a quadratic cost function to minimize the lateral distance and orientation deviation from a given reference path and to punish high curvature rates for comfort.

In the case that trajectory optimization is infeasible using the selected lateral or longitudinal driving corridor, we select a driving corridor with the next highest cumulative drivable area for optimization until either a fail-safe trajectory is identified or no further driving corridors remain. In the rare event that no feasible fail-safe trajectory is found, the previously verified trajectory is further executed.

Guarantees of our verification technique. To comply with legal safety, autonomous vehicles must not collide with any legal behaviour of other traffic participants:

$$\forall t \geq t_0 : occ(x(t)) \cap (\mathcal{L}^e(t) \cup \mathcal{Q}) = \emptyset \tag{2}$$

where the operator $occ(x)$ relates the state x of the autonomous vehicle to the set of occupied points in the position domain as $occ(x) : \mathcal{X} \rightarrow Pow(\mathbb{R}^n)$, where $Pow(\mathbb{R}^n)$ is the power set of $\mathbb{R}^n$.

Using the principle of induction, we sketch the proof that our technique ensures legal safety according to equation (2). For the base case ($c = 1$), for $t \geq t_0$, the autonomous vehicle is initially in a safe state in which it can remain. Only if I_c can be successfully verified will the autonomous vehicle start executing $I_c^{\text{safe}}||F_c$ from t_c. This trajectory is collision-free at all discrete time steps $t'_k \in [t_c, t_c + \Delta_c^{\text{safe}} + T_{F_c}]$ against all legal behaviours $\mathcal{L}(t) \supseteq \mathcal{L}^e(t)$ of other traffic participants and the area $\mathcal{Q}$ outside the road. If no new intended trajectory can be successfully verified in a subsequent verification cycle before $t_c + \Delta_c^{\text{safe}} + T_{F_c}$, the fail-safe trajectory F_c transitions the autonomous vehicle to a standstill in a safe terminal state at $t_c + \Delta_c^{\text{safe}} + T_{F_c}$, which is legally safe for all future times. For the inductive step, assuming that the verification result of cycle $c = r$, for any $r \in \mathbb{N}_+$, ensures legal safety, we show that legal safety is also ensured regardless of the verification result of cycle $c + 1$. If the verification is unsuccessful, the autonomous vehicle continues to execute the trajectory $I_{c-i}^{\text{safe}}||F_{c-i}, i \in \{0, \ldots, c-1\}$ of the previous cycle c that ensures legal safety by definition. If the verification is successful in cycle $c + 1$, the autonomous vehicle executes $I_{c+1}^{\text{safe}}||F_{c+1}$ from t_{c+1}. In this case, we can apply the same reasoning as in the base case to demonstrate that legal safety is also ensured from t_{c+1} with the verified trajectory $I_{c+1}^{\text{safe}}||F_{c+1}$.

To ensure that the autonomous vehicle is collision-free along I^{safe} and F in continuous time and despite control disturbances and model uncertainties, we refer to the approach in ref. [67].

Data availability

All data gathered and reported in this study are available in the Supplementary data file. This includes the environment model, the intended trajectory and the verification result of each verification cycle for all scenarios.

Code availability

The code to visualize and analyse the gathered data and obtained results of this study are included in the Supplementary data file.

Received: 30 October 2019; Accepted: 4 August 2020;
Published online: 14 September 2020

References

1. Favarò, F., Eurich, S. & Nader, N. Autonomous vehicles' disengagements: trends, triggers and regulatory limitations. *Accid. Anal. Prev.* **110**, 136–148 (2018).
2. Anderson, J. M. et al. *Autonomous Vehicle Technology*: *A Guide for Policymakers* (Rand Corporation, 2016).
3. Koopman, P. & Wagner, M. Autonomous vehicle safety: an interdisciplinary challenge. *IEEE Intell. Transportation Syst. Mag.* **9**, 90–96 (2017).
4. Kalra, N. & Paddock, S. M. Driving to safety: how many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Res. A Policy Practice* **94**, 182–193 (2016).
5. Seshia, S. A., Sadigh, D. & Sastry, S. S. Towards verified artificial intelligence. Preprint at https://arxiv.org/abs/1606.08514 (2017).
6. Schwarting, W., Alonso-Mora, J. & Rus, D. Planning and decision-making for autonomous vehicles. *Annu. Rev. Control Robot. Autonomous Syst.* **1**, 187–210 (2018).
7. United Nations Economic Commission for Europe. *Convention on Road Traffic. United Nations Conference on Road Traffic* (United Nations, 1968); consolidated version of 2006.
8. Vanholme, B., Gruyer, D., Lusetti, B., Glaser, S. & Mammar, S. Highly automated driving on highways based on legal safety. *IEEE Trans. Intell. Transportation Syst.* **14**, 333–347 (2013).
9. Althoff, M. & Dolan, J. M. Online verification of automated road vehicles using reachability analysis. *IEEE Trans. Robotics* **30**, 903–918 (2014).
10. Koopman, P. & Wagner, M. Challenges in autonomous vehicle testing and validation. *SAE Int. J. Transportation Safety* **4**, 15–24 (2016).
11. Dahl, J., de Campos, G. R., Olsson, C. & Fredriksson, J. Collision avoidance: a literature review on threat-assessment techniques. *IEEE Trans. Intell. Vehicles* **4**, 101–113 (2019).
12. Tumova, J., Hall, G. C., Karaman, S., Frazzoli, E. & Rus, D. Least-violating control strategy synthesis with safety rules. In *Proceedings of the 16th International Conference on Hybrid Systems*: *Computation and Control* 1–10 (HSCC, 2013).
13. Kress-Gazit, H., Fainekos, G. E. & Pappas, G. J. Temporal-logic-based reactive mission and motion planning. *IEEE Trans. Robotics* **25**, 1370–1381 (2009).
14. Fraichard, T. & Asama, H. Inevitable collision states—a step towards safer robots? In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* 388–393 (IEEE, 2003).
15. Chan, N., Kuffner, J. & Zucker, M. Improved motion planning speed and safety using regions of inevitable collision. In *17th CISM-IFToMM Symposium on Robot Design, Dynamics and Control* 103–114 (Springer, 2008).
16. Koller, T., Berkenkamp, F., Turchetta, M. & Krause, A. Learning-based model predictive control for safe exploration. In *Proceedings of the 2018 IEEE International Conference on Decision and Control* 6059–6066 (IEEE, 2018).

17. Wabersich, K. P. & Zeilinger, M. N. Linear model predictive safety certification for learning-based control. In *Proceedings of the IEEE International Conference on Decision and Control* 7130–7135 (IEEE, 2018).
18. Sadraddini, S. & Belta, C. A provably correct MPC approach to safety control of urban traffic networks. In *Proceedings of the American Control Conference* 1679–1684 (2016).
19. Ames, A. D. et al. Control barrier functions: theory and applications. In *Proceedings of the 18th European Control Conference* 3420–3431 (IEEE, 2019).
20. Tedrake, R., Manchester, I. R., Tobenkin, M. & Roberts, J. W. LQR-trees: feedback motion planning via sums-of-squares verification. *Int. J. Robotics Res.* **29**, 1038–1052 (2010).
21. Li, W., Sadigh, D., Sastry, S. S. & Seshia, S. A. Synthesis for human-in-the-loop control systems. In *Proceedings of the International Conference on Tools and Algorithms for the Construction and Analysis of Systems* 470–484 (Springer, 2014).
22. Jalalmaab, M., Fidan, B., Jeon, S. & Falcone, P. Guaranteeing persistent feasibility of model predictive motion planning for autonomous vehicles. In *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium* 843–848 (IEEE, 2017).
23. Danielson, C., Weiss, A., Berntorp, K. & Di Cairano, S. Path planning using positive invariant sets. In *Proceedings of the 55th International Conference on Decision and Control* 5986–5991 (IEEE, 2016).
24. Herbert, S. L. et al. FaSTrack: a modular framework for fast and guaranteed safe motion planning. In *Proceedings of the 56th International Conference on Decision and Control* 1517–1522 (IEEE, 2017).
25. Falcone, P., Ali, M. & Sjöberg, J. Predictive threat assessment via reachability analysis and set invariance theory. *IEEE Trans. Intell. Transportation Syst.* **12**, 1352–1361 (2011).
26. Vaskov, S. et al. Towards provably not-at-fault control of autonomous robots in arbitrary dynamic environments. In *Proc. Robotics: Science and Systems* 1–9 (2019).
27. Lefèvre, S., Vasquez, D. & Laugier, C. A survey on motion prediction and risk assessment for intelligent vehicles. *ROBOMECH J.* **1**, 1–14 (2014).
28. Gindele, T., Brechtel, S. & Dillmann, R. Learning driver behavior models from traffic observations for decision making and planning. *IEEE Intell. Transportation Syst. Mag.* **7**, 69–79 (2015).
29. Bahram, M., Hubmann, C., Lawitzky, A., Aeberhard, M. & Wollherr, D. A combined model- and learning-based framework for interaction-aware maneuver prediction. *IEEE Trans. Intell. Transportation Syst.* **17**, 1538–1550 (2016).
30. Deo, N., Rangesh, A. & Trivedi, M. M. How would surround vehicles move? A unified framework for maneuver classification and motion prediction. *IEEE Trans. Intell. Vehicles* **3**, 129–140 (2018).
31. Ghahramani, Z. Probabilistic machine learning and artificial intelligence. *Nature* **521**, 452–459 (2015).
32. Tang, C., Chen, J. & Tomizuka, M. Adaptive probabilistic vehicle trajectory prediction through physically feasible Bayesian recurrent neural network. In *Proceedings of the 2019 IEEE International Conference on Robotics and Automation* 3846–3852 (IEEE, 2019).
33. Pool, E. A. I., Kooij, J. F. P. & Gavrila, D. M. Context-based cyclist path prediction using recurrent neural networks. In *Proceedings of the 2019 IEEE Intelligent Vehicles Symposium* 824–830 (IEEE, 2019).
34. Wu, A. & How, J. Guaranteed infinite horizon avoidance of unpredictable, dynamically constrained obstacles. *Autonomous Robots* **32**, 227–242 (2012).
35. Bouraine, S., Fraichard, T. & Salhi, H. Provably safe navigation for mobile robots with limited field-of-views in dynamic environments. *Autonomous Robots* **32**, 267–283 (2012).
36. Yang, Y., Zhang, J., Cai, K. & Prandini, M. Multi-aircraft conflict detection and resolution based on probabilistic reach sets. *IEEE Trans. Control Syst. Technol.* **25**, 309–316 (2017).
37. Nager, Y., Censi, A. & Frazzoli, E. What lies in the shadows? Safe and computation-aware motion planning for autonomous vehicles using intent-aware dynamic shadow regions. In *Proceedings of the 2019 IEEE International Conference on Robotics and Automation* 5800–5806 (IEEE, 2019).
38. McNaughton, M., Urmson, C., Dolan, J. M. & Lee, J.-W. Motion planning for autonomous driving with a conformal spatiotemporal lattice. In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation* 4889–4895 (IEEE, 2011).
39. Werling, M., Kammel, S., Ziegler, J. & Gröll, L. Optimal trajectories for time-critical street scenarios using discretized terminal manifolds. *Int. J. Robotics Res.* **31**, 346–359 (2012).
40. Zucker, M. et al. CHOMP: covariant Hamiltonian optimization for motion planning. *Int. J. Robotics Res.* **32**, 1164–1193 (2013).
41. Ziegler, J., Bender, P., Dang, T. & Stiller, C. Trajectory planning for Bertha—a local, continuous method. In *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium* 450–457 (IEEE, 2014).
42. Hult, R., Zanon, M., Gros, S. & Falcone, P. An MIQP-based heuristic for optimal coordination of vehicles at intersections. In *Proceedings of the 2018 IEEE International Conference on Decision and Control* 2783–2790 (IEEE, 2018).
43. Sun, Z., Hsu, D., Jiang, T., Kurniawati, H. & Reif, J. H. Narrow passage sampling for probabilistic roadmap planning. *IEEE Trans. Robotics* **21**, 1105–1115 (2005).
44. LaValle, S. M. in *Planning Algorithms* 79–80 (Cambridge Univ. Press, 2006).
45. Schouwenaars, T., De Moor, B., Feron, E. & How, J. Mixed integer programming for multi-vehicle path planning. In *Proceedings of the 2001 European Control Conference* 2603–2608 (IEEE, 2001).
46. Qian, X., Altché, F., Bender, P., Stiller, C. & de La Fortelle, A. Optimal trajectory planning for autonomous driving integrating logical constraints: an MIQP perspective. In *Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems* 205–210 (IEEE, 2016).
47. Park, J., Karumanchi, S. & Iagnemma, K. Homotopy-based divide-and-conquer strategy for optimal trajectory planning via mixed-integer programming. *IEEE Trans. Robotics* **31**, 1101–1115 (2015).
48. Gutjahr, B., Gröll, L. & Werling, M. Lateral vehicle trajectory optimization using constrained linear time-varying MPC. *IEEE Trans. Intell. Transportation Syst.* **18**, 1586–1595 (2016).
49. Zhan, W., Chen, J., Chan, C.-Y., Liu, C. & Tomizuka, M. Spatially-partitioned environmental representation and planning architecture for on-road autonomous driving. In *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium* 632–639 (IEEE, 2017).
50. Mohy-ud-Din, H. & Muhammad, A. Detecting narrow passages in configuration spaces via spectra of probabilistic roadmaps. In *Proceedings of the 2010 ACM Symposium on Applied Computing* 1294–1298 (ACM, 2010).
51. Do, Q. H., Mita, S. & Yoneda, K. Narrow passage path planning using fast marching method and support vector machine. In *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium* 630–635 (IEEE, 2014).
52. Bender, P., Taş, Ö. S., Ziegler, J. & Stiller, C. The combinatorial aspect of motion planning: maneuver variants in structured environments. In *Proceedings of the 2015 IEEE Intelligent Vehicles Symposium* 1386–1392 (IEEE, 2015).
53. Archer, J. & Vogel, K. *The Traffic Safety Problems in Urban Areas. Technical Report* (KTH Stockholm, 2000).
54. Shalev-Shwartz, S., Shammah, S. & Shashua, A. On a formal model of safe and scalable self-driving cars. Preprint at https://arxiv.org/pdf/1708.06374.pdf (2018).
55. Liebenwein, L. et al. Compositional and contract-based verification for autonomous driving on road networks. In *Robotics Research, Springer Proceedings in Advanced Robotics* Vol. 10, 163–181 (Springer, 2020).
56. Trautman, P. & Krause, A. Unfreezing the robot: navigation in dense, interacting crowds. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems* 797–803 (IEEE, 2010).
57. Menéndez-Romero, C., Winkler, F., Dornhege, C. & Burgard, W. Maneuver planning for highly automated vehicles. In *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium* 1458–1464 (IEEE, 2017).
58. Althoff, M. & Magdici, S. Set-based prediction of traffic participants on arbitrary road networks. *IEEE Trans. Intell. Vehicles* **1**, 187–202 (2016).
59. Koschi, M. & Althoff, M. SPOT: a tool for set-based prediction of traffic participants. In *Proceedings of the 2017 IEEE Intelligent Vehicles Symposium* 1686–1693 (IEEE, 2017).
60. Koschi, M., Pek, C., Beikirch, M. & Althoff, M. Set-based prediction of pedestrians in urban environments considering formalized traffic rules. In *Proceedings of the 21st International Conference on Intelligent Transportation Systems* 2704–2711 (IEEE, 2018).
61. Pek, C. & Althoff, M. Computationally efficient fail-safe trajectory planning for self-driving vehicles using convex optimization. In *Proceedings of the 2018 IEEE International Conference on Intelligent Transportation Systems* 1447–1454 (IEEE, 2018).
62. Manzinger, S., Pek, C. & Althoff, M. Using reachable sets for trajectory planning of automated vehicles. *IEEE Trans. Intell. Vehicles* https://doi.org/10.1109/TIV.2020.3017342 (2020).
63. Paden, B., Čáp, M., Yong, S. Z., Yershov, D. & Frazzoli, E. A survey of motion planning and control techniques for self-driving urban vehicles. *IEEE Trans. Intell. Vehicles* **1**, 33–55 (2016).
64. González, D., Pérez, J., Milanés, V. & Nashashibi, F. A review of motion planning techniques for automated vehicles. *IEEE Trans. Intell. Transportation Syst.* **17**, 1135–1145 (2016).
65. Magdici, S., Ye, Z. & Althoff, M. Determining the maximum time horizon for vehicles to safely follow a trajectory. In *Proceedings of the 20th International Conference on Intelligent Transportation Systems* 1893–1899 (IEEE, 2017).
66. Héry, E., Masi, S., Xu, P. & Bonnifait, P. Map-based curvilinear coordinates for autonomous vehicles. In *Proceedings of the 20th International Conference on Intelligent Transportation Systems* 1–7 (IEEE, 2017).
67. Schürmann, B. et al. Ensuring drivability of planned motions using formal methods. In *Proceedings of the 20th International Conference on Intelligent Transportation Systems* 1661–1668 (IEEE, 2017).

Acknowledgements

We thank S. Kaster for his support in implementing the prediction and S. Steyer for providing the object detection and tracking algorithms. We also thank C. Schürmann for the voice-overs in the Supplementary Videos. This work was partially supported by the BMW Group within the CAR@TUM project, the German Federal Ministry of Economics and Technology through the research initiative Ko-HAF, and the German Research Foundation (DFG) under grants AL 1185/4-2 and AL 1185/3-2.

Author contributions

C.P., S.M. and M.K. developed the verification technique during replanning. M.K. developed the concept and algorithms for the set-based prediction. C.P. and S.M. developed the concept and algorithms for the drivable area computation, driving corridor identification and fail-safe trajectory planning. M.A. developed the main concept of online verification by integrating set-based prediction and fail-safe trajectory generation. He also developed the underlying algorithms for reachability analysis and led the research project. C.P., S.M. and M.K. designed and conducted the experiments and collected the data. The Article and the Supplementary Information were written by C.P., S.M. and M.K.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s42256-020-0225-y.

Correspondence and requests for materials should be addressed to C.P., S.M. or M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.