

# Spieltheoretischer prädiktiver Regler für Interaktion im gemischten Verkehr

Mohamed-Khalil Bouzidi\* und Ehsan Hashemi†

**Zusammenfassung:** Dieser Artikel präsentiert eine integrierte Trajektorienplanung und Regelung für Verkehrsszenarien in denen Interaktion mit Fahrzeugen, die von Menschen gefahren werden, erforderlich ist. Für eine sichere Bewegungsplanung berücksichtigt die vorgeschlagene Methode die Interaktion zwischen dem automatisierten Fahrzeug und anderen Fahrzeugen anhand der Spieltheorie. Das Framework umfasst ein neuartiges inverses Differentialspiel basierend auf einer LSTM Architektur, um die Zielfunktion des menschlichen Fahrers online zu schätzen. Diese Schätzung wird dann von einem spieltheoretischen prädiktiven Regler genutzt, um die Trajektorie des menschlich gesteuerten Fahrzeugs zu präzidieren und das autonome Fahrzeug zu steuern. Das entwickelte System wird in einer Human-in-the-Loop Probandenstudie unter Verwendung der CarSim High-Fidelity-Simulation evaluiert.

**Schlüsselwörter:** Long Short-Term Memory (LSTM), Mensch-Maschine-Interaktion, Spieltheoretischer Modellprädiktiver Regler, Trajektorienplanung

## 1 Einleitung

Testfahrten autonomer Fahrzeuge(ADS) haben in den letzten Jahren gezeigt, dass diese Systeme bei bestimmten Herausforderungen an ihre Grenzen stoßen. Eines dieser Probleme, das so genannte Frozen Robot Problem(FRP) [1], wurde bereits mehrfach beobachtet. Insbesondere in komplexen Verkehrsszenarien, wie z.B. bei Einfädelungsmanövern, scheitern autonome Fahrzeuge bei der Durchführung. Dieses Fehlverhalten lässt sich durch die Entkopplung der Prädiktionsaufgabe (die die Trajektorien der Verkehrsteilnehmer in der Szene prädiziert) von der Planungsaufgabe (die eine sichere Trajektorie auf der Basis einer der vorgeschalteten Prädiktionen generiert) erklären: Die umliegenden Fahrzeuge werden als dynamische Hindernisse betrachtet, die nicht auf die gefahrene Trajektorie des ADS reagieren, wodurch durchführbare Manöver nicht erkannt werden. Um in der Lage zu sein, die Situation korrekt zu antizipieren, muss das ADS wissen, dass die umliegenden Fahrzeuge mit dem ADS interagieren, um Kollisionen zu vermeiden. Dies erfordert gekoppelte Vorhersage- und Planungsaufgaben, da die Fahrer ihr Verhalten an die Trajektorien der anderen Fahrzeuge anpassen und umgekehrt.

In dieser Hinsicht bietet die Spieltheorie einen geeigneten Rahmen, um diese gleichzeitige Planung und Prädiktion anzugehen. Da im gemischten Verkehr oft Verkehrsteilnehmer nicht unbedingt kooperativ agieren, werden solche Interaktionen oft als unkooperatives Spiel z.B. als

---

\*Mohamed-Khalil Bouzidi ist bei Continental AG tätig und Doktorand an der Freien Universität Berlin, Deutschland. Die Arbeit wurde im Rahmen seiner Masterarbeit am NODE Lab, Fachbereich Maschinenbau, University of Alberta, Edmonton, AB, Kanada, durchgeführt. (e-mail: m.bouzidi@fu-berlin.de).

†Ehsan Hashemi ist im Fachbereich Maschinenbau, University of Alberta, Edmonton, AB, Kanada, (e-mail: ehashemi@ualberta.ca).

Nash-Equilibrium (NE) modelliert. Existierende Ansätze für Trajektorienplanung [2,3] beruhen jedoch auf der Annahme, dass die Zielfunktion des Fahrers dem ADS bekannt ist.

Die Fahrer haben in der Regel unterschiedliche Ziele und Fahrstile, was eine Online-Schätzung der Zielfunktion des Menschen erforderlich macht. Zur Schätzung von Zielfunktionen wird in der Literatur das inverse Differentialspiel (IDG) verwendet. Bestehende Methoden sind jedoch entweder zu rechenintensiv, um sie online auszuführen, wie z. B. direkte Methoden [4], oder inverse Reinforcement Learning basierte Methoden [5] oder sie beruhen auf zu starken Annahmen, wie Methoden, die von Inverse Optimal Control [6] abgeleitet sind.

In [7] wird ein Framework vorgeschlagen, die ein Feed-Forward-Netz für die Online-IDG verwendet, und ein weiteres Netz, das im nächsten Schritt eine NE findet. Das Fehlen eines Fahrzeugmodells und die begrenzten (weniger als 50) Trainingsdaten bieten jedoch keine Garantie für fahrbare Trajektorien. In [4] wird ein unscented Kalman-Filter (UKF) zur Online-Schätzung der Zielfunktionsparameter verwendet (d. h. eine direkte Methode für IDG) welches dem spieltheoretische modellprädiktive Regler (GT MPC) [3] als Input dient. Die Methoden verwendet das kinematische Einspurmodell, das das Fahrzeugverhalten bei Autobahn-Einfädelmanövern, die in der Regel mit hoher Geschwindigkeit durchgeführt werden und eine ungleichmäßige Drehmomentverteilung an jeder Achse aufweisen können, nicht angemessen approximieren kann. Dies kann zur Generierung unrealisierbarer Trajektorien führen. Ein einfaches Austauschen des Modells ist allerdings nicht möglich, da eine höhere Anzahl von Zuständen und Parametern der Zielfunktion (aufgrund komplexerer dynamischer Modelle) zu rechnerischen Herausforderungen für den vorgeschlagenen UKF-basierten Schätzer führt.

Um diese Probleme anzugehen, schlagen wir ein interaktionsbewusstes Framework basierend auf der Spieltheorie vor, der das dynamische Einspurmodell verwendet, um dynamisch realisierbare Trajektorien zu generieren, und ein LSTM-Netzwerk für die Online-Schätzung des Verhaltens von Fahrzeugen bei Einfädelungen in Echtzeit verwendet. Unser Framework wurde in Einfädelungsszenarien im gemischten Verkehr evaluiert (mit menschlichen Probanden in einem Human-in-the-Loop-Setup und unter Verwendung hochgenauer *CarSim*-Simulationen), was aufgrund der zuvor genannten Gründe für andere spieltheoretische Ansätze [2, 4, 7] nicht möglich ist. Unsere Hauptbeiträge lassen sich daher wie folgt zusammenfassen:

- Entwicklung eines LSTM-basierten inversen Differentialspiels zur Online-Schätzung der Zielfunktions des menschlichen Fahrers für eine prädiktive Regelungsstrategie.
- Entwurf eines Unknown Input Observer zur Schätzung der Steuergrößen des umgebenden Fahrzeugs mithilfe von Bord-Sensorikmessungen.
- Entwicklung einer spieltheoretischen modellprädiktiven Regelung, die die laterale Fahrzeugdynamik berücksichtigt, um laterale Stabilität und dynamische Realisierbarkeit der geplanten Trajektorien für eine sicherere Mensch-ADS-Interaktion zu gewährleisten.

## 2 Problem Definition und Modellbeschreibung

Betrachtet wird ein autonomes Ego-Fahrzeug welches in die Autobahn einfädeln möchte. Ein von einem Menschen gesteuertes Fahrzeug (auf der linken Seite) interagiert entsprechend mit dem ADS. Das Ziel besteht darin, eine kollisionsfreie Trajektorie zu planen, dem das Ego-Fahrzeug folgen kann. Um die interaktive Natur des Autobahn-Einfädelns zu erfassen, modellieren wir dieses Szenario als ein unkooperatives Spiel. Wir wählen das NE als Lösungskonzept,

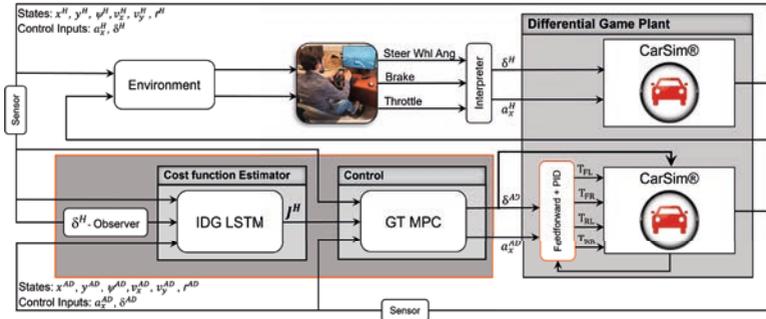


Abbildung 1: Das interaktionsbewusste spieltheoretische Trajektorienplanungs- und Regelung Framework unter Verwendung eines Unknown Input Observer und LSTM

da bei Nash-Spielen kein Spieler einen strukturellen Vorteil gegenüber anderen Spielern hat, im Gegensatz zur Leader-Follower-Struktur eines Stackelberg-Spiels.

**Definition 1.** Das Nash-Gleichgewicht ist ein Lösungskonzept, das entsteht, wenn jeder Spieler/Agent  $i$  gleichzeitig und optimal bezüglich seiner eigenen Zielfunktion  $J^i$  handelt und seinen Annahmen über die Strategien der anderen Spieler  $j$ , wobei diese Annahmen korrekt sind. Dies entspricht:

$$J^i(x^*, u_i^*, u_j^*) < J^i(x^*, u_i, u_j^*) \quad \forall i. \quad (1)$$

Spieltheoretisch ausgedrückt, kann dieses Autobahn-Einfädungsszenario als eine Szene beschrieben werden, in der zwei Agenten  $i \in AD, H$  (automatisiertes Fahrsystem und von einem Menschen gesteuertes Fahrzeug) in einem dynamischen Spiel mit endlichem Horizont verhandeln. Um diese Szene als Nash-Spiel zu modellieren, treffen wir folgende Annahme:

**Annahme 1.** Der Mensch handelt rational, um seine Zielfunktion  $J^H$  zu minimieren.

Tatsächlich wird das Entscheidungsverhalten und die Reaktion des Menschen bei Fahrmanövern oft in der Literatur [8] durch einen optimalen Regler modelliert, da Menschen das Fahrverhalten anderer basierend auf ihrem Wissen vorhersagen und darauf rational reagieren.

Es sollte beachtet werden, dass es keine Kommunikation zwischen den Fahrzeugen wie V2V gibt. Daher hat das ADS keinen Zugriff auf die Zielfunktion des anderen Fahrzeugs (d.h. Agent  $H$ ). Allerdings wird diese Zielfunktion benötigt, um das NE zu berechnen. Daher wird die Zielfunktion online mittels eines IDG und der beobachteten Trajektorien der Fahrzeuge geschätzt. Mit der Schätzung der Zielfunktion wird die Lösung des NE basierend auf dem Modell des Systems berechnet, wie in Abbildung 1 veranschaulicht.

Wir modellieren das Spiel als Differentialspiel, da die Fahrzeugkinematik und -dynamik naturgemäß kontinuierlich sind. Das Differentialspiel wird durch ein dynamisches System beschrieben:  $\dot{x}(t) = f(x(t), u^H(t), u^A(t))$ . Aufgrund von Situationen mit hoher Schlupfgefahr beim Einfädeln verwenden wir ein dynamisches Einspurmodell, um Trajektorien präzisieren, die die Manövrierfähigkeiten (und Grenzwerte der Fahrzeugführung) korrekt einschätzen und somit dynamisch machbare Trajektorien generieren. Daher werden die seitlichen Reifenkräfte

$F_{y,f}$  und  $F_{y,r}$  berücksichtigt, wobei von linearen Reifenkräften ausgegangen wird. Der Schlupfwinkel an der Vorder- und Hinterachse des Fahrzeugs wird jeweils durch  $\alpha_f = \delta - \frac{l_f r + v_y}{v_x}$  und  $\alpha_r = \frac{l_r r - v_y}{v_x}$  definiert. Hierbei steht  $\delta$  für den Lenkwinkel der Vorderachse und die Längsgeschwindigkeit, Quergeschwindigkeit und Giergeschwindigkeit des Fahrzeugs im Fahrzeugbezugssystem (bezogen auf den Fahrzeugschwerpunkt, CG) werden durch  $v_x$ ,  $v_y$  und  $r$  bezeichnet. Die Fahrzeugdynamik kann dann wie folgt beschrieben werden:

$$\dot{v}_x = v_y r + a_x, \quad \dot{v}_y = -v_x r + \left( \frac{F_{y,f} + F_{y,r}}{m} \right), \quad \dot{r} = \ddot{\psi} = \left( \frac{F_{y,f} l_f - F_{y,r} l_r}{I_z} \right) \quad (2)$$

wo  $a_x$  der Längsbeschleunigung entspricht,  $m$  ist die Fahrzeugmasse und die Abstände der Vorder- und Hinterachse zum Schwerpunkt des Fahrzeugs werden mit  $l_f$  und  $l_r$  bezeichnet. Die Dynamik der Längs-/Querposition ergibt sich zu  $[\dot{x}, \dot{y}]^\top = R_\psi [v_x, v_y]^\top$  mit der Rotationsmatrix  $R_\psi$  sowohl für das ADS als auch für das von einem Menschen gesteuerte Fahrzeug. Diese Modelle werden dann gestapelt, mit Ausnahme des ersten Zustands (d.h., der Längsposition, bei der der inter-vehikuläre Abstand  $e_x$  als Zustandsvariable verwendet wird), um das dynamische Modell für das Differentialspiel mit den Steuergrößen  $\mathbf{u}^i = [a_x^i, \delta^i]^\top$  für  $i \in AD, H$  und den erweiterten Zuständen  $\mathbf{x} = [e_x, y^{AD}, y^H, \psi^{AD}, \psi^H, v_x^{AD}, v_x^H, v_y^{AD}, v_y^H, r^{AD}, r^H]^\top$  zu bilden.

Um die Zielfunktion des menschlichen Fahrers  $H$  zu bestimmen, verwenden wir einen Basisfunktion-Ansatz. Es wird angenommen, dass die Zielfunktion beider Agenten  $i \in AD, H$  durch  $J^i = \int \theta^i \phi^i(t) dt$  beschrieben werden kann. Die Parameter  $\theta^H$  müssen mit Hilfe eines IDG identifiziert werden. Die Parameter  $\theta^{AD}$  des ADS sind einstellbar. Wir beschreiben das Ziel des Fahrzeugs  $H$  durch (3), wobei der Term zur Minimierung der Quergeschwindigkeit und somit des Seitenschlupfwinkels entscheidend ist, um die laterale Stabilität sicherzustellen.

$$J^i = \int \theta_1^i (v_x^i - v_{x,d}^i)^2 + \theta_2^i (y^i - y_d^i)^2 + \theta_3^i (v_y^i)^2 + \theta_4^i \tanh \left( \underbrace{\sqrt{\left( e_x^2 + \left( \frac{b}{a} (y^H - y^{AD}) \right)^2 \right)}}_{\Gamma} \right) - a \right) + \theta_5^i (a_x^i)^2 + \theta_6^i (\delta^i)^2 dt \quad \forall i \quad (3)$$

Die Zielfunktion (3) kann in  $J^i = \int \theta^i \phi^i(t) dt$  mit den Parametern  $\theta^i = [\theta_1^i, v_{x,d}^i, \theta_2^i, y_d^i, \theta_3^i, \theta_4^i, \theta_5^i, \theta_6^i]^\top$  und der Basisfunktion  $\phi^i(t) = [(v_x^i)^2, -2v_x^i, (y^i)^2, -2y^i, (v_y^i)^2, \tanh(\Gamma), (a_x^i)^2, (\delta^i)^2]^\top$  transformiert werden.

### 3 Unknown Input Observer Design

Das ADS ist mit Kameras, LiDARs und Radarsensoren zur Detektion von Verkehrsteilnehmern und zur Messung ihrer relativen Abstände/Ausrichtungen ausgestattet. Der Lenkwinkel des von einem Menschen gesteuerten Fahrzeugs ist nicht messbar. Daher wird ein Unknown Input Observer (UIO) [9] entworfen, um diesen Eingang mithilfe des linearen lateralen Fahrzeugdynamikmodells  $\dot{\zeta} = A\zeta + B\mathbf{u}$  mit den Zuständen  $\zeta = [v_y^H, r^H]^\top$ , messbaren Ausgängen  $\mathbf{y} = \zeta$  und dem unbekanntem Eingang  $\mathbf{u} = \delta^H$  abzuschätzen. Der folgende Beobachter approximiert den Lenkeingang des umgebenden Fahrzeugs, wobei  $\mathbf{z}$  die Zustandsvariable für den Beobachter ist.

$$\hat{\delta}^H = [B^\top B]^{-1} B^\top (\dot{\mathbf{z}} - \hat{\zeta} - A\hat{\mathbf{y}}), \quad \dot{\mathbf{z}} = F\mathbf{z}(t) + K\mathbf{y}(t), \quad \hat{\zeta} = \mathbf{z}(t) + \bar{H}\mathbf{y}(t). \quad (4)$$

Die Fehlerdynamik des Schätzerfehler  $e = \zeta - \hat{\zeta}$  entspricht dementsprechend:

$$\dot{e} = (A - \bar{H}A - K_1)e + [F - (A - \bar{H}A - K_1)]z + [K_2 - (A - \bar{H}A - K_1)]y + (\bar{H} - I)Bu, \quad (5)$$

wobei die asymptotische Stabilität der Fehlerdynamik erreicht wird, wenn die Matrizen des Beobachter wie folgt entworfen werden:  $\bar{H} = B[B^T B]^{-1}B^T$ ,  $F = (I - \bar{H}C)A - K_1C$ ,  $K_2 = F\bar{H}$ , wobei  $K_1$  so gewählt wird, dass  $F$  Hurwitz ist [9]. Einfädelungsszenarien beinhalten in der Regel Geschwindigkeitsvariationen aufgrund Interaktion mit dem umgebenden Fahrzeug, weshalb die Matrizen unter Verwendung der gemessenen Geschwindigkeit adaptiert werden.

## 4 Spieltheoretischer Prädiktiver Regler

Unter Verwendung des nichtlinearen Systemmodells  $\dot{x} = f(x, u^H, u^{AD})$  und der Zielfunktionen  $J^{AD}$ ,  $J^H$  entwerfen wir ein spieltheoretisches Modellprädiktiver Regler (GT MPC), der das Open-Loop-Problem für einen gegebenen Prädiktionshorizont löst, aber die Anfangszustandswerte  $x_0$  in jedem Zeitschritt mit dem tatsächlich beobachteten Zustand aktualisiert.

Die Lösung für das Open-Loop-NE (OLNE) kann numerisch bestimmt werden, indem es in ein nichtlineares Programm (NLP) mit den Variablen  $\mathbf{X} = [x_1, x_2, \dots, x_{N_p}]^T, \mathbf{U}^i = [u_1^i, \dots, u_{N_c}^i]^T$  umgewandelt wird, wobei  $N_p$  den Vorhersagehorizont und  $N_c$  den Steuerungshorizont angibt. Wir formulieren das Lagrange-Multiplikator-Problem wie folgt:

$$L^i = J^i + \sum_{k=0}^{N_p} (\lambda_k^i)^T [x_{k+1} - f_k(x_k, u_k^H, u_k^{AD})] \quad \forall i \quad (6)$$

Entsprechend müssen für ein OLNE die folgenden Optimalitätsbedingungen erfüllt sein:

$$\begin{aligned} \mathbf{G}^i &= \nabla_{\mathbf{X}, \mathbf{U}^i} L^i(\mathbf{X}, \mathbf{U}^{AD}, \mathbf{U}^H, \boldsymbol{\lambda}^{AD}, \boldsymbol{\lambda}^H) = \mathbf{0} \quad \forall i \\ \mathbf{C} &= [x_1 - f_k(x_0, u_0) \dots x_{N_p} - f_k(x_{N_p-1}, u_{N_p-1})]^T = \mathbf{0} \end{aligned} \quad (7)$$

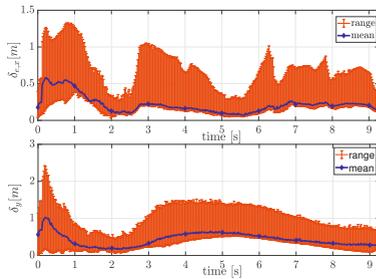
Um eine Lösung zu finden, die den Optimalitätsbedingungen genügt, verwenden wir CasADi [10] mit folgender Suchrichtung:

$$\begin{aligned} p_{ND} &= -\mathbf{H}^{-1} \mathbf{G} \\ \mathbf{G} &= [\mathbf{G}^{AD}, \mathbf{G}^H, \mathbf{C}]^T, \quad \mathbf{H} = \nabla_{(\mathbf{X}, \mathbf{U}, \boldsymbol{\lambda})} \mathbf{G}(\mathbf{X}, \mathbf{U}^{AD}, \mathbf{U}^H, \boldsymbol{\lambda}^{AD}, \boldsymbol{\lambda}^H) \end{aligned} \quad (8)$$

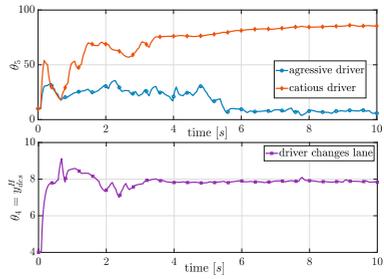
## 5 Inverses Differentialspiel

Die IDG dient hier dazu mithilfe von zuvor beobachteten Trajektorien  $x(t)$ ,  $u^H(t)$  und  $u^{AD}(t)$  die Parameter  $\theta^H$  der Zielfunktion des menschlichen Fahrers  $J = \int \theta^H \phi^H(t) dt$  zu bestimmen.

Die LSTM-Struktur eignet sich gut zur Schätzung von Zuständen aus sukzessiv eintreffenden Features, wobei die Länge der Input-Zeitreihe variieren kann. Daher muss die Dauer des Manövers nicht im Voraus bekannt sein. Das Ziel besteht darin, den Parametervektor  $\theta_k^H = \theta^H$  zu schätzen (d. h., ein Sequence-to-One-Regressionsproblem). Das Input Layer erhält den aktuellen Zustand  $x_k$ , die Steuergroße des menschlichen Fahrers  $u_k^H$  sowie die vorherige Steuergroße des ADS  $u_{k-1}^{AD}$ . Dies liegt daran, dass die Entscheidung des Systems nicht vorhergesagt werden kann, indem nur das menschliche Fahrzeug isoliert beobachtet wird, da die Interaktion seine Entscheidung beeinflusst.



a Evaluation des Prädiktionsfehlers (longitudinale Distanz und laterale Position) aller Probanden



b Schätzung der Zielfunktion bei aggressiven und vorsichtigen Fahrer und bei Fahrspurwechsel

**Anmerkung 1.** Bei Nash-Spielen ist bekannt, dass jeder Spieler seine Basisfunktion minimieren möchte, um die Zielfunktion mit positiven Gewichten der Basisfunktionen zu minimieren. Zu jedem Zeitschritt wird die Basisfunktion  $\phi_k^H t_k$  berechnet und als Input für das neuronale Netz verwendet. Dieser zusätzliche Input dient dazu Vorwissen einzubeziehen und die Generalisierungsleistung des Netzes zu verbessern.

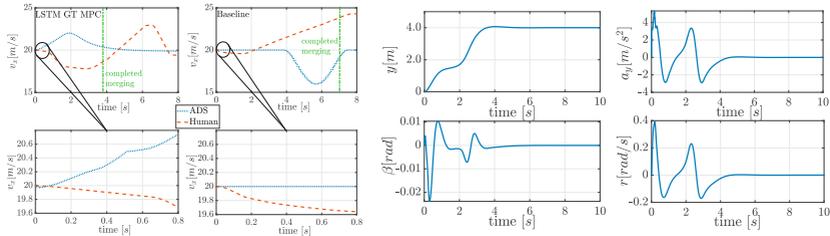
Das neuronale Netz dient als kontinuierliche Abbildung mit nichtlinearen Dynamiken mit  $\theta^H = \mathcal{N}(xk, \mathbf{u}_k^H, \mathbf{u}_{k-1}^{AD}, \phi_k^H)$ . Es besteht aus zwei LSTM-Layer mit jeweils 80 und 60 hidden units. Dahinter werden zwei fully connected Layer mit 40 Neuronen und der RELU-Aktivierungsfunktion geschaltet.

Ein GT MPC wird verwendet, um zufällige Trajektorien synthetisch für das Training zu generieren, wobei angenommen wird, dass die von GT MPC generierten Trajektorien ausreichend konsistent mit realen Fahrscenarien unter Beteiligung von Menschen sind. Dies wurde durch umfangreiche Tests im letzten Abschnitt bestätigt. Dementsprechend werden unterschiedliche Längen von Trajektorien für das Training verwendet, so dass das LSTM-Modell den Parametervektor schätzen kann, nachdem es einen Teil der Trajektorie beobachtet hat.

## 6 Evaluation

Die Evaluation unseres Ansatzes wird anhand einer Probandenstudie mit 6 verschiedenen Personen durchgeführt, die nicht mit dem Algorithmus vertraut sind. Das Ziel der Studie ist es, zu analysieren, ob die spieltheoretische Modellierung für diesen Anwendungsfall gerechtfertigt ist (d.h. ob Annahme 1 gerechtfertigt ist) und ob das Framework robust gegenüber unterschiedlichen Fahrstilen ist. Für den Fahrsimulator verwenden wir highfidelity Simulationen in *CarSim*. Die Probanden nutzen zweimal das Human-in-the-Loop-Setup. Die erste Hälfte der Teilnehmer beginnt mit einem konventionellen Planer [11], die zweite Hälfte mit dem neuen Planer. Für das autonome Fahrzeug haben wir ein anspruchsvolles Szenario gewählt, bei dem beide Fahrzeuge mit derselben Anfangsgeschwindigkeit von  $v_{x,0} = 20$ , m/s und einem Längsabstand von  $e_{x,0} = 1$ , m starten.

In allen Fällen war unser Ansatz in der Lage, ein sicheres Einfädeln durchzuführen. Das Verhältnis des erfolgreichen Einfädelns von vorne betrug 4/6. Die durchschnittliche Einfädelzeit betrug 4.26 Sekunden. Der Baseline-Planer konnte das Einfädelmanöver in 5/6 der Fälle erfolgreich abschließen, und das Verhältnis des erfolgreichen Einfädelns von vorne betrug 3/6.



a Vergleich des neuen Ansatzes mit dem Baseline-Planer an einem ausgewählten Beispiel      b Laterales Verhalten des Autonomen Fahrzeugs während des ausgewählten Beispiels

Die durchschnittliche Einfädelzeit betrug 7.18 Sekunden. Die durchschnittliche Einfädelzeit und das erfolgreiche Einfädeln von vorne sind wichtige Indikatoren dafür, wie erfolgreich das autonome Fahrzeug in Einfädel-Szenarien sein wird, da sich hinter dem aktuellen menschlichen Fahrzeug ein weiteres menschliches Fahrzeug befinden könnte, mit dem das autonome Fahrsystem erneut interagieren muss. Die Ergebnisse zeigen daher, dass unser Planer weniger anfällig für das Frozen Robot Problem ist.

Die Ursache hierfür lässt sich anhand eines Beispiels (siehe Abb. 3a) erklären. Hier bremst der menschliche Fahrer leicht ab, um dem autonomen Fahrzeug das Einfädeln zu signalisieren. Das entwickelte interaktionsbewusste Framework antizipiert die Reaktion des anderen Fahrers für ein sicheres Einfädeln, d.h. es kann vorhersehen, dass der Mensch leicht bremst, weil er mit dem autonomen System interagiert, und in Zukunft stärker bremsen wird. Dieses Vorwissen führt zu einer verbesserten Vorhersage der Trajektorien (siehe Abb. 2a), wobei auch zwischen aggressiven und vorsichtigen Fahrern, unter anderem anhand des Wertes der Kollisionsvermeidungsgewichtung  $\theta_5$ , unterschieden werden kann (siehe Abb. 2b). So weiß das autonome Fahrzeug früher, wann es beschleunigen (oder bremsen) sollte, und kann sicher auf die Autobahn einfädeln, wobei die laterale Stabilität trotz der Abweichung zwischen dem angenommenen Einfachspurmodell und *CarSim* aufrechterhalten wird (z.B. siehe Abb. 3b). Auf der anderen Seite wartet der Baseline-Planer in diesem Beispiel darauf, dass der Mensch stärker bremst, weil er nicht erkennt, dass der Mensch auf das autonome System reagiert. Als Ergebnis reagiert der Baseline-Regler nicht, das menschlich gesteuerte Fahrzeug beschleunigt und das autonome Fahrzeug muss hinter dem Fahrzeug einfädeln. Dieses Verhalten führt zum Frozen Robot Problem im dichten Verkehr, wenn die nachfolgenden Fahrzeuge ähnlich handeln.

## 7 Fazit

Ein interaktionsbewusstes Framework für Trajektorienplanung und -regelung wurde entwickelt, der in der Lage ist, mit einem vom Menschen gesteuerten Fahrzeug zu interagieren. Es werden Einfädelungsszenarien betrachtet, welche wir als Nash-Differentialspiel mit Hilfe des dynamischen Einspurmodells beschreiben. Basierend darauf wurde ein spieltheoretische prädiktiver Regler entwickelt und getestet. Der prädiktive Regler nimmt die Zielfunktion des menschlichen Fahrers des anderen Fahrzeugs als Input, die in Echtzeit mit Hilfe eines lernbasierten inversen Differentialspiels geschätzt wird. Die experimentelle Auswertung in *CarSim* mit einem Human-

in-the-Loop-Setup hat gezeigt, dass die spieltheoretische Modellierung solcher Szenarien vorteilhaft für diese Art von Szenarien und der interaktionsbewusste Regelungssystem in der Lage ist, verschiedene Fahrstile der von Menschen gesteuerten (d.h. umgebenden) Fahrzeuge mit guter Recheneffizienz zu handhaben.

## Literatur

- [1] P. Trautman and A. Krause, “Unfreezing the robot: Navigation in dense, interacting crowds,” in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 797–803.
- [2] D. Fridovich-Keil, E. Ratner, L. Peters, A. Dragan, and C. Tomlin, “Efficient iterative linear-quadratic approximations for nonlinear multi-player general-sum differential games,” 05 2020, pp. 1475–1481.
- [3] S. L. Cleac'h, M. Schwager, and Z. Manchester, “ALGAMES: A fast solver for constrained dynamic games,” in *Robotics: Science and Systems XVI*. Robotics: Science and Systems Foundation, 2020.
- [4] S. Le Cleacâh, M. Schwager, and Z. Manchester, “Lucidgames: Online unscented inverse dynamic games for adaptive trajectory prediction and planning,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5485–5492, 2021.
- [5] N. Mehr, M. Wang, and M. Schwager, “Maximum-entropy multi-agent dynamic games: Forward and inverse solutions,” 2021.
- [6] J. Inga, *Inverse Dynamic Game Methods for Identification of Cooperative System Behavior*. Karlsruher Institut für Technologie, 2020.
- [7] P. Geiger and C.-N. Straehle, “Learning game-theoretic models of multiagent trajectories using implicit layers,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35.
- [8] C. C. MacAdam, “Understanding and modeling the human driver,” *Vehicle System Dynamics*, vol. 40, pp. 101 – 134, 2003.
- [9] M. Darouach, M. Zasadzinski, and S. J. Xu, “Full-order observers for linear systems with unknown inputs,” *IEEE Transactions on Automatic Control*, vol. 39, no. 3, pp. 606–609, 1994.
- [10] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, “CasADi – A software framework for nonlinear optimization and optimal control,” *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.
- [11] Y. Ding, W. Zhuang, L. Wang, J. Liu, L. Guvenc, and Z. Li, “Safe and optimal lane-change path planning for automated driving,” *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 235, no. 4, pp. 1070–1083, 2021.