Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 56-2 (2023) 8320–8326

Validation of Stochastic Optimal Control Models for Goal-Directed Human Movements on the Example of Human Driving Behavior

Philipp Karg * **Simon Stoll** * **Simon Rothfuß** * **Sören Hohmann** *

* *Institute of Control Systems (IRS), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany (e-mail: philipp.karg@kit.edu)*

Abstract: Stochastic Optimal Control models represent the state-of-the-art in modeling goal-directed human movements. The linear-quadratic sensorimotor (LQS) model based on signal-dependent noise processes in state and output equation is the current main representative. With our newly introduced Inverse Stochastic Optimal Control algorithm building upon two bi-level optimizations, we can identify its unknown model parameters, namely cost function matrices and scaling parameters of the noise processes, for the first time. In this paper, we use this algorithm to identify the parameters of a deterministic linear-quadratic, a linear-quadratic Gaussian and a LQS model from human measurement data to compare the models' capability in describing goal-directed human movements. Human steering behavior in a simplified driving task shown to posses similar features as point-ot-point human hand reaching movements serves as our example movement. The results show that the identified LQS model outperforms the others with statistical significance. Particularly, the average human steering behavior is modeled significantly better by the LQS model. This validates the positive impact of signal-dependent noise processes on modeling human average behavior.

Copyright © 2023 The Authors. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: Human Movement Modeling, Human Driving Behavior, Inverse Optimal Control, Inverse Stochastic Optimal Control, Stochastic Optimal Control

1. INTRODUCTION

Stochastic Optimal Control (SOC) models are the state-of-the-art approach to describe human movements to a single goal (Gallivan et al. (2018)). Since they model basic optimality principles underlying human movements, like minimum intervention (Todorov and Jordan (2002)), they are able to model the average behavior and the variability patterns of human limb movements. SOC models, and closed-loop optimization approaches in general, challenge traditional open-loop optimization models consisting of a separation between movement planning and execution (see e.g. Uno et al. (1989); Flash and Hogan (1985)). Moreover, due to their white-box character SOC models show higher generalizability and transferability between tasks than black-box models (alias behavioral cloning, see e.g. Ross and Bagnell (2010)). The current main SOC model is the linear-quadratic (LQ) sensorimotor (LQS) model, not least due to its computational tractability. The LQS model builds upon the well-known LQ Gaussian (LQG) framework. To this model, a control-dependent noise process is added to the state equation describing that faster movements are performed more inaccurately. Furthermore, a state-dependent noise process is added to the output equation describing the higher inaccuracy of the perception of faster movements. In order to validate SOC models in specific movement tasks, an identification method is needed to determine the unknown model parameters, namely relative weights of cost function candidates and scaling parameters of the noise processes, from human measurement data. Closely associated with the validation, such an identification method enables the analysis of optimality principles underlying human movements in a specific task as well (Franklin and Wolpert (2011); Todorov (2004)), i.e. preferred cost function candidates, influence of different noise processes on average and variability patterns as well as comparison to adapted models with omitted or added model parts. However, until very recently (Karg et al. (2022)), such an identification method that identifies all unknown parameters of the LQS model from human measurement data was non-existent. The identified models and especially the conclusions drawn from the analysis of underlying human optimality principles are the crucial starting point for the design of human-machine systems that closely interact with the human in a cooperative, intuitive and safe manner: prediction, imitation and classification of human behavior is enabled as well as a model-based automation design.

Kolekar et al. (2018) motivate that human steering behavior is comparable to point-to-point reaching, widely considered as movement task in the neuroscientific community when human hand movements are studied. Since, in addition, a close cooperation between human and automation in steering a vehicle shows great potential to increase safety, for example in takeover scenarios (see Karg

2405-8963 Copyright © 2023 The Authors. This is an open access article under the CC BY-NC-ND license.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2023.10.1021

et al. (2021)), human steering behavior in a simplified driving task acts as our example movement to investigate SOC models for goal-directed human movements. Describing human driving behavior via an optimal control framework traces back to (MacAdam (1981)) but only in recent years these optimal control models were extended to include neuroscientific findings (Kolekar et al. (2018); Nash and Cole (2015); Markkula (2014); Sentouh et al. (2009)). Nash and Cole (2015) propose the first stochastic human driver model that is based on the LQG framework. However, their model lacks signal-dependent noise processes which are promising to be considered (Kolekar et al. (2018); Markkula (2014)). Kolekar et al. (2018) make a first step to identify and validate the LQS model for human steering behavior but do not provide a rigor and generally applicable identification method that determines all model parameters. Moreover, they focus on describing the variability patterns of the human driver and hence, a comparison between the LQS, LQG and a deterministic LQ model regarding human average steering behavior is missing.

In this work, we want to deepen this line of research. First, we clarify that our steering task is comparable to a point-to-point reaching task by showing that the kinematic features of the movements in both tasks match. This strengthens the motivation done by Kolekar et al. (2018) that steering is similar to reaching. Then, we use our newly introduced Inverse Stochastic Optimal Control (ISOC) algorithm (Karg et al. (2022)) to identify all parameters of the LQS model from human measurement data and compare its modeling performance to reduced versions of it, i.e. to a LQG and LQ model. Here, our comparison and analysis is particularly focused on human average behavior and the influence of signal-dependent noise processes on it. This influence has theoretically been proven in our previous work (Karg et al. (2022)).

2. MATERIALS AND METHODS

In the following section the LQS model and our ISOC algorithm (Karg et al. (2022)) are introduced. Afterwards, the experimental setup is explained.

2.1 Stochastic Optimal Control Models for Goal-Directed Human Movements

The human biomechanics and the system the human interacts with, e.g. a vehicle with steering wheel, is modeled via the linear time-discrete system dynamics

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{\Sigma}^{\boldsymbol{\xi}}\boldsymbol{\alpha}_t + \sum_{i=1}^{c} \varepsilon_t^{(i)} \boldsymbol{C}_i \boldsymbol{u}_t \tag{1}$$

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{x}_t + \boldsymbol{\Sigma}^{\boldsymbol{\omega}}\boldsymbol{\beta}_t + \sum_{i=1}^{d} \epsilon_t^{(i)} \boldsymbol{D}_i \boldsymbol{x}_t, \tag{2}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ denotes the system state, $\boldsymbol{u} \in \mathbb{R}^m$ the control variable (typically neural activation of muscles), $\boldsymbol{y} \in \mathbb{R}^r$ the human observation and $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{H}$ matrices of appropriate dimension that may depend on time. When the additional time index t is used, e.g. $\boldsymbol{x}_t$, the stochastic process of the corresponding variable is denoted. Furthermore, $\boldsymbol{\alpha}_t \in \mathbb{R}^p$ and $\boldsymbol{\varepsilon}_t = \left[\varepsilon_t^{(1)} \ \dots \ \varepsilon_t^{(c)}\right]^\top$ are standard white Gaussian noise processes that constitute the additive and control-dependent noise process of state equation (1) with their scaling matrices $\boldsymbol{\Sigma}^{\boldsymbol{\xi}}$ and $\boldsymbol{C}_i = \sigma_i^{\boldsymbol{u}} \boldsymbol{B}\boldsymbol{F}_i$ of appropriate dimension. In (2), $\boldsymbol{\beta}_t \in \mathbb{R}^q$ and $\boldsymbol{\epsilon}_t = \left[\epsilon_t^{(1)} \ \dots \ \epsilon_t^{(d)}\right]^\top$ are standard white Gaussian noise processes forming the additive and state-dependent noise process of the output equation with their corresponding scaling matrices $\boldsymbol{\Sigma}^{\boldsymbol{\omega}}$ and $\boldsymbol{D}_i = \sigma_i^{\boldsymbol{x}} \boldsymbol{H}\boldsymbol{G}_i$ of appropriate dimension. The stochastic processes $\boldsymbol{\alpha}_t$, $\boldsymbol{\varepsilon}_t$, $\boldsymbol{\beta}_t$ and $\boldsymbol{\epsilon}_t$ are independent to each other and to $\boldsymbol{x}_t$. The goal of the human, e.g. achievement of a desired steering angle under minimal effort, is modeled by the performance criterion

$$J = \mathrm{E}\left\{ \boldsymbol{x}_N^\top \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{t=0}^{N-1} \boldsymbol{x}_t^\top \boldsymbol{Q} \boldsymbol{x}_t + \boldsymbol{u}_t^\top \boldsymbol{R} \boldsymbol{u}_t \right\}, \tag{3}$$

where $\boldsymbol{Q}_N$, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are symmetric and of appropriate dimension. $\boldsymbol{Q}_N$, $\boldsymbol{Q}$ are positive semi-definite and $\boldsymbol{R}$ is positive definite. It is assumed that the human solves the LQS control problem consisting of the minimization of (3) with respect to (1) and (2) by an admissible control strategy. An approximate (suboptimal) solution to the LQS control problem is given by Todorov (2005) with $\boldsymbol{u}_t = -\boldsymbol{L}_t \hat{\boldsymbol{x}}_t$ and $\hat{\boldsymbol{x}}_{t+1} = \boldsymbol{A}\hat{\boldsymbol{x}}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{K}_t(\boldsymbol{y}_t - \boldsymbol{H}\hat{\boldsymbol{x}}_t) + \boldsymbol{\eta}_t$, where $\boldsymbol{L}_t$ and $\boldsymbol{K}_t$ are computed by iterating between

$$\boldsymbol{L}_t = \Big(\boldsymbol{R} + \boldsymbol{B}^\top \boldsymbol{Z}_{t+1}^{\boldsymbol{x}} \boldsymbol{B} + \sum_i \boldsymbol{C}_i^\top \left(\boldsymbol{Z}_{t+1}^{\boldsymbol{x}} + \boldsymbol{Z}_{t+1}^{\boldsymbol{e}} \right) \boldsymbol{C}_i \Big)^{-1} \boldsymbol{B}^\top \boldsymbol{Z}_{t+1}^{\boldsymbol{x}} \boldsymbol{A} \tag{4}$$

and

$$\boldsymbol{K}_t = \boldsymbol{A}\boldsymbol{P}_t^{\boldsymbol{e}} \boldsymbol{H}^\top \Big(\boldsymbol{H}\boldsymbol{P}_t^{\boldsymbol{e}} \boldsymbol{H}^\top + \boldsymbol{\Omega}^{\boldsymbol{\omega}} + \sum_i \boldsymbol{D}_i \left(\boldsymbol{P}_t^{\boldsymbol{e}} + \boldsymbol{P}_t^{\hat{\boldsymbol{x}}} + \boldsymbol{P}_t^{\hat{\boldsymbol{x}}\boldsymbol{e}} + \boldsymbol{P}_t^{\boldsymbol{e}\hat{\boldsymbol{x}}} \right) \boldsymbol{D}_i^\top \Big)^{-1}. \tag{5}$$

In (4) and (5), $\boldsymbol{Z}_t^{\boldsymbol{e}}$, $\boldsymbol{Z}_t^{\boldsymbol{x}}$, $\boldsymbol{P}_t^{\boldsymbol{e}}$, $\boldsymbol{P}_t^{\hat{\boldsymbol{x}}}$, $\boldsymbol{P}_t^{\boldsymbol{e}\hat{\boldsymbol{x}}}$ and $\boldsymbol{P}_t^{\hat{\boldsymbol{x}}\boldsymbol{e}}$ follow from recursive equations (see Todorov (2005)).

Suppose the cost function matrices $\boldsymbol{Q}_N$, $\boldsymbol{Q}$, $\boldsymbol{R}$ and the noise scaling parameters $\boldsymbol{\Sigma}^{\boldsymbol{\xi}}$, $\sigma_i^{\boldsymbol{u}}$, $\boldsymbol{\Sigma}^{\boldsymbol{\omega}}$, $\sigma_i^{\boldsymbol{x}}$ are known, the average human behavior (mean $\mathrm{E}\{\boldsymbol{x}_t\}$) and the human variability pattern (covariance $\boldsymbol{\Omega}_t^{\boldsymbol{x}} = \mathrm{cov}(\boldsymbol{x}_t, \boldsymbol{x}_t) = \mathrm{cov}(\boldsymbol{x}_t)$) predicted by the LQS model are given by Lemma 1.

Lemma 1. Let the LQS control problem be defined by the minimization of (3) with respect to (1) and (2). Let the (approximate) solution be given by $\boldsymbol{L}_t$ (4) and $\boldsymbol{K}_t$ (5). Then, the mean $\mathrm{E}\{\boldsymbol{x}_t\}$ and covariance $\boldsymbol{\Omega}_t^{\boldsymbol{x}}$ of $\boldsymbol{x}_t$ are

$$\begin{bmatrix} \mathrm{E}\{\boldsymbol{x}_{t+1}\} \\ \mathrm{E}\{\hat{\boldsymbol{x}}_{t+1}\} \end{bmatrix} = \boldsymbol{\mathcal{A}}_t \begin{bmatrix} \mathrm{E}\{\boldsymbol{x}_t\} \\ \mathrm{E}\{\hat{\boldsymbol{x}}_t\} \end{bmatrix}, \tag{6}$$

$$\begin{bmatrix} \boldsymbol{\Omega}_{t+1}^{\boldsymbol{x}} & \boldsymbol{\Omega}_{t+1}^{\boldsymbol{x}\hat{\boldsymbol{x}}} \\ \boldsymbol{\Omega}_{t+1}^{\hat{\boldsymbol{x}}\boldsymbol{x}} & \boldsymbol{\Omega}_{t+1}^{\hat{\boldsymbol{x}}} \end{bmatrix} = \boldsymbol{\mathcal{A}}_t \begin{bmatrix} \boldsymbol{\Omega}_t^{\boldsymbol{x}} & \boldsymbol{\Omega}_t^{\boldsymbol{x}\hat{\boldsymbol{x}}} \\ \boldsymbol{\Omega}_t^{\hat{\boldsymbol{x}}\boldsymbol{x}} & \boldsymbol{\Omega}_t^{\hat{\boldsymbol{x}}} \end{bmatrix} \boldsymbol{\mathcal{A}}_t^\top + \begin{bmatrix} \boldsymbol{\Omega}^{\boldsymbol{\xi}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega}^{\boldsymbol{\eta}} + \boldsymbol{K}_t \boldsymbol{\Omega}^{\boldsymbol{\omega}} \boldsymbol{K}_t^\top \end{bmatrix} + \begin{bmatrix} \bar{\boldsymbol{\Omega}}_t^{\hat{\boldsymbol{x}}} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{\Omega}}_t^{\boldsymbol{x}} \end{bmatrix} \tag{7}$$

with $\bar{\boldsymbol{\Omega}}_t^{\hat{\boldsymbol{x}}} = \sum_i \boldsymbol{C}_i \boldsymbol{L}_t \left(\boldsymbol{\Omega}_t^{\hat{\boldsymbol{x}}} + \mathrm{E}\{\hat{\boldsymbol{x}}_t\} \mathrm{E}\{\hat{\boldsymbol{x}}_t\}^\top \right) \boldsymbol{L}_t^\top \boldsymbol{C}_i^\top$, $\bar{\boldsymbol{\Omega}}_t^{\boldsymbol{x}} = \sum_i \boldsymbol{K}_t \boldsymbol{D}_i \left(\boldsymbol{\Omega}_t^{\boldsymbol{x}} + \mathrm{E}\{\boldsymbol{x}_t\} \mathrm{E}\{\boldsymbol{x}_t\}^\top \right) \boldsymbol{D}_i^\top \boldsymbol{K}_t^\top$,

$$\boldsymbol{\mathcal{A}}_t = \begin{bmatrix} \boldsymbol{A} & -\boldsymbol{B}\boldsymbol{L}_t \\ \boldsymbol{K}_t\boldsymbol{H} & \boldsymbol{A} - \boldsymbol{K}_t\boldsymbol{H} - \boldsymbol{B}\boldsymbol{L}_t \end{bmatrix} \tag{8}$$

and initial values $\mathrm{E}\{\hat{\boldsymbol{x}}_0\} = \hat{\boldsymbol{x}}_0 = \mathrm{E}\{\boldsymbol{x}_0\}$ and

$$\begin{bmatrix} \boldsymbol{\Omega}_0^{\boldsymbol{x}} & \boldsymbol{\Omega}_0^{\boldsymbol{x}\hat{\boldsymbol{x}}} \\ \boldsymbol{\Omega}_0^{\hat{\boldsymbol{x}}\boldsymbol{x}} & \boldsymbol{\Omega}_0^{\hat{\boldsymbol{x}}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Omega}_0^{\boldsymbol{x}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}. \tag{9}$$

Proof. See (Karg et al., 2022, Lemma 2).

Corollary 1. The mean $\mathrm{E}\{\boldsymbol{x}_t\}$ of the closed-loop LQS system depends on the noise scaling parameters $\sigma_i^{\boldsymbol{u}}$ in the fully observable case ($\hat{\boldsymbol{x}}_t = \boldsymbol{x}_t$) and on $\sigma_i^{\boldsymbol{u}}$, $\sigma_i^{\boldsymbol{x}}$, $\boldsymbol{\Sigma}^{\boldsymbol{\xi}}$ and $\boldsymbol{\Sigma}^{\boldsymbol{\omega}}$ in the partially observable case.

Proof. From (6) and (8) the dependency of $\mathrm{E}\{\boldsymbol{x}_t\}$ on $\boldsymbol{L}_t$ (4) follows. In the fully observable case, $\boldsymbol{Z}_t^{\boldsymbol{e}}$ drops out in (4) (see Todorov (2005)) but $\boldsymbol{L}_t$ still depends on $\boldsymbol{C}_i$ and thus $\sigma_i^{\boldsymbol{u}}$. In the partially observable case, the additional dependencies follow from the dependency of $\boldsymbol{L}_t$ on $\boldsymbol{Z}_t^{\boldsymbol{e}}$ which in turn depends on $\boldsymbol{K}_t$ (see Todorov (2005)).

Corollary 1 shows that the human average behavior $\mathrm{E}\{\boldsymbol{x}_t\}$ predicted by the LQS model depends not only on the cost function matrices but also on the scaling parameters of the noise processes. Hence, they could act as a crucial part of the LQS model improving the performance in describing human average behavior.

2.2 Inverse Stochastic Optimal Control

Regarding the LQS model introduced in the previous section, the cost function matrices $\boldsymbol{Q}_N$, $\boldsymbol{Q}$, $\boldsymbol{R}$ and the noise (scaling) parameters $\boldsymbol{\Sigma}^{\boldsymbol{\xi}}$, $\sigma_i^{\boldsymbol{u}}$, $\boldsymbol{\Sigma}^{\boldsymbol{\omega}}$, $\sigma_i^{\boldsymbol{x}}$ are unknown and need to be identified when looking at human movements in practice. To simplify notation, we combine the cost function parameters in a vector $\boldsymbol{s} \in \mathbb{R}^S$ and the noise parameters in a vector $\boldsymbol{\sigma} \in \mathbb{R}^\Sigma$ (see Karg et al. (2022)). Now, we assume that our measurement data consists of a set $\{\boldsymbol{M}\boldsymbol{x}_t^{*,(k)}\}$ ($k \in \{1, \ldots, K\}$) of time-discrete trajectories $\boldsymbol{M}\boldsymbol{x}_t^{*,(k)}$ where $\boldsymbol{x}_t^{*,(k)}$ is a realization of the stochastic process $\boldsymbol{x}_t^*$. It results in the closed-loop LQS system with the unknown cost function $\boldsymbol{s}^*$ and noise parameters $\boldsymbol{\sigma}^*$ of the human. With $\boldsymbol{M} \in \mathbb{R}^{\bar{n} \times n}$, we consider that not all system states are measured, i.e. $\boldsymbol{M}$ follows from the identity matrix by deleting rows corresponding to not measured states. The identification problem is finally formalized in Problem 1 [1].

Problem 1. Let estimates $\hat{\boldsymbol{m}}_t \approx \mathrm{E}\{\boldsymbol{M}\boldsymbol{x}_t^*\}$ and $\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*} \approx \boldsymbol{M}\boldsymbol{\Omega}_t^{\boldsymbol{x}^*}\boldsymbol{M}^\top$ computed from the measurement data be given. Find parameters $\tilde{\boldsymbol{s}}$ and $\tilde{\boldsymbol{\sigma}}$ such that they lead to $\tilde{\boldsymbol{x}}_t$ in the closed-loop LQS system with $\mathrm{E}\{\boldsymbol{M}\tilde{\boldsymbol{x}}_t\} = \mathrm{E}\{\boldsymbol{M}\boldsymbol{x}_t^*\}$ and $\boldsymbol{M}\boldsymbol{\Omega}_t^{\tilde{\boldsymbol{x}}}\boldsymbol{M}^\top = \boldsymbol{M}\boldsymbol{\Omega}_t^{\boldsymbol{x}^*}\boldsymbol{M}^\top$.

In order to solve Problem 1, we have introduced an ISOC algorithm based on two bi-level optimizations in our previous work (Karg et al. (2022)). We formulate a direct optimization problem by defining a performance criterion J_{ISOC} describing how well the mean $\mathrm{E}\{\boldsymbol{x}_t\}$ and covariance $\boldsymbol{\Omega}_t^{\boldsymbol{x}}$ predicted by the LQS model with parameters $\boldsymbol{s}$ and $\boldsymbol{\sigma}$ match the measurement data $\hat{\boldsymbol{m}}_t$ and $\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$. The objective function J_{ISOC} is chosen based on the Variance Accounted For (VAF) metric:

$$J_{\mathrm{ISOC}} = \frac{\boldsymbol{w}_{\mathrm{m}}^\top \boldsymbol{m}^{\mathrm{VAF}} + \boldsymbol{w}_{\mathrm{v}}^\top \operatorname{vec}\left(\boldsymbol{\Omega}^{\mathrm{VAF}}\right)}{\|\boldsymbol{w}_{\mathrm{m}}\|_1 + \|\boldsymbol{w}_{\mathrm{v}}\|_1}, \tag{10}$$

[1] The problem extends typical Inverse Optimal Control or Inverse Reinforcement Learning problems by additionally searching for scaling parameters of noise processes.

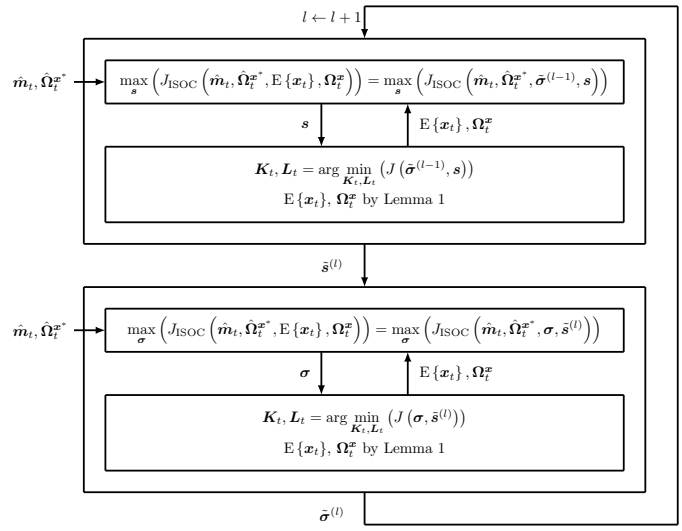

Fig. 1. Bi-level-based ISOC algorithm.

where for the elements m_i^{VAF} and $\Omega_{ij}^{\mathrm{VAF}}$

$$m_i^{\mathrm{VAF}} = \left(1 - \frac{\sum_{t=0}^N \left(\left(\mathrm{E}\{\boldsymbol{M}\boldsymbol{x}_t\}\right)_i - \hat{m}_{i,t}\right)^2}{\sum_{t=0}^N \left(\hat{m}_{i,t} - \frac{1}{N+1}\sum_t \hat{m}_{i,t}\right)^2}\right), \tag{11}$$

$$\Omega_{ij}^{\mathrm{VAF}} = \left(1 - \frac{\sum_{t=0}^N \left(\left(\boldsymbol{M}\boldsymbol{\Omega}_t^{\boldsymbol{x}}\boldsymbol{M}^{\mathrm{T}}\right)_{ij} - \hat{\Omega}_{ij,t}^{\boldsymbol{x}^*}\right)^2}{\sum_{t=0}^N \left(\hat{\Omega}_{ij,t}^{\boldsymbol{x}^*} - \frac{1}{N+1}\sum_t \hat{\Omega}_{ij,t}^{\boldsymbol{x}^*}\right)^2}\right) \tag{12}$$

holds. In (10), $\boldsymbol{w}_{\mathrm{m}} \in \mathbb{R}^{\bar{n}}$ and $\boldsymbol{w}_{\mathrm{v}} \in \mathbb{R}^{\bar{n}\bar{n}}$ denote weighting vectors. Due to the VAF metric $J_{\mathrm{ISOC}} \in (-\infty, 1]$ holds with 1 corresponding to a perfect fit between predicted and measured data. Therefore, the direct ISOC optimization problem consists of maximizing J_{ISOC}: $\max_{\boldsymbol{s},\boldsymbol{\sigma}} \left(J_{\mathrm{ISOC}}\left(\hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}, \mathrm{E}\{\boldsymbol{x}_t\}, \boldsymbol{\Omega}_t^{\boldsymbol{x}}\right)\right)$. The evaluation of J_{ISOC} with specific values for $\boldsymbol{s}$ and $\boldsymbol{\sigma}$ depends on calculating the model predictions $\mathrm{E}\{\boldsymbol{x}_t\}$ and $\boldsymbol{\Omega}_t^{\boldsymbol{x}}$ which in turn depends on solving the forward optimal control problem, i.e. computing $\boldsymbol{K}_t$ and $\boldsymbol{L}_t$. Hence, maximizing J_{ISOC} leads to a bi-level optimization problem where the upper level optimization consists of maximizing J_{ISOC} and the lower level of calculating $\boldsymbol{K}_t$, $\boldsymbol{L}_t$, $\mathrm{E}\{\boldsymbol{x}_t\}$ and $\boldsymbol{\Omega}_t^{\boldsymbol{x}}$. In order to keep this optimization computationally tractable, we apply an alternating optimization approach by computing the best estimate $\tilde{\boldsymbol{s}}^{(l)}$ for a given estimate $\tilde{\boldsymbol{\sigma}}^{(l-1)}$ in the first step of one iteration l and in the second vice versa. This leads to the iteration between two bi-level optimizations shown in Fig. 1. Furthermore, in the lower levels of both bi-level optimizations our newly introduced Lemma 1 is utilized to simplify the computation of $\mathrm{E}\{\boldsymbol{x}_t\}$ and $\boldsymbol{\Omega}_t^{\boldsymbol{x}}$ by applying recursive calculations only. Finally, for solving the upper level optimizations the parameters $\boldsymbol{s}$ and $\boldsymbol{\sigma}$ are divided into parameter sets where each parameter is supposed to be in at least one set and where the parameters in one set are assumed to have a mutually correlated influence on J_{ISOC}. Then, grid searches with fixed grid size are performed on these parameter sets iteratively until no further improvement with the current grid size is achieved. Then, the grid size is shrinked. Further details regarding the ISOC algorithm can be found in (Karg et al. (2022)).

Fig. 2. Screen visualization of the simplified steering task. The study participants interacted with an active steering wheel to change the horizontal position of the squared marker and were asked to follow the piecewise constant reference trajectory as best as possible.

2.3 Human Driving Experiment

The human measurement data was gathered by a study where 14 participants (12 male and 2 female subjects) aged between 19 and 29 (24.9 in average) performed a simplified steering task. The participants interacted with an active steering wheel which was equipped with an incremental encoder (40000 steps per full rotation) to measure the steering angle φ with a sampling frequency of 100 Hz. Based on the steering angle measurements torque was applied by the steering wheel such that it mimicked the dynamics of a spring-damper system. With the steering wheel the participants were able to move a marker (square) on a screen on a horizontal line with fixed vertical position (cf. Figure 2). The horizontal position of the marker was a linear mapping of the steering angle. On the display, a reference trajectory for the marker (cf. Figure 2) was shown and the participants were asked to follow it as best as possible. The reference trajectory moved downwards with constant velocity. The reference trajectory shown for the marker is directly connected to a reference trajectory $\varphi_{\mathrm{r},t}$ for the steering angle. The trial of one participant started with a familiarization phase of 210 s consisting of different segments with a steering angle reference with constant curvature where the curvature varied between those segments. Afterwards, the reference consisted of 14 repetitions of the same piecewise constant pattern: $\varphi_{\mathrm{r}} = 0$, $\varphi_{\mathrm{r}} = +\frac{2}{3}\pi$, $\varphi_{\mathrm{r}} = 0$ and $\varphi_{\mathrm{r}} = -\frac{2}{3}\pi$, each for a duration of 100 time steps.

Regarding the data preparation, we first applied a cubic spline smoothing on the measured steering angle to reduce measurement noise. Afterwards, we calculate the steering angle velocity $\dot{\varphi}_t$ via numerical differentiation. Since the measured data consists of 14 repetitions of the $+\frac{2}{3}\pi$-step as well as 14 repetitions of the $-\frac{2}{3}\pi$-step, we calculate estimates of human average behavior ($\hat{\boldsymbol{m}}_t$) and human variability pattern ($\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$) for each of these movements from the corresponding 14 repetitions. Hereto, we determine the starting time for each single movement repetition as the last point in time where $\varphi \approx 0$ and $\dot{\varphi} < 0.1\,\mathrm{s}^{-1}$ holds. In some segments with $\varphi_{\mathrm{r}} = 0$ the steering angle velocity is always above this threshold. Then, we increase the threshold in $0.01\,\mathrm{s}^{-1}$-steps until a maximum value of $0.4\,\mathrm{s}^{-1}$. If even this maximally relaxed condition is not fulfilled, we remove this single movement repetition from the further analysis since the human clearly starts not from an equilibrium condition in this case. Finally, the remaining repetitions are used to calculate the mean ($\hat{\boldsymbol{m}}_t$) and covariance estimate ($\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$) of φ_t and $\dot{\varphi}_t$. The movement duration is defined as the averaged movement duration of those repetitions which fulfill the condition for a movement ending ($\varphi \approx \pm\frac{2}{3}\pi$ and $\dot{\varphi} < 0.1\,\mathrm{s}^{-1}$). This data preparation is done for every subject.

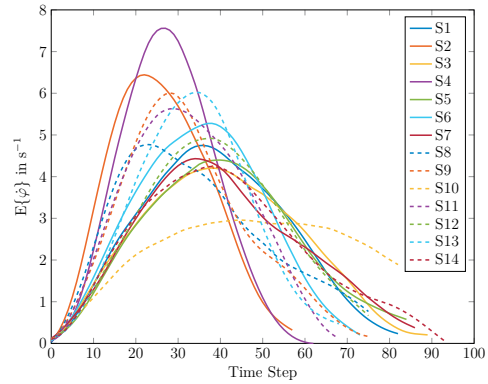

Fig. 3. Steering angle velocity profiles of all 14 subjects (S1-S14) for the $+\frac{2}{3}\pi$-step.

Fig. 3 shows the steering angle velocity profiles for the $+\frac{2}{3}\pi$-step achieved by the averaging procedure described before. Remarkably, the profiles possess very similar kinematic features as point-to-point reaching movements: they are all single-peaked and predominantly bell-shaped and nearly symmetric. The majority of the profiles is furthermore slightly left-skewed which indicates that the subjects tend to perform fast movements (Engelbrecht (2001)). Moreover, the maximum to mean ratios of $\mathrm{E}\{\dot{\varphi}\}$ range from 1.78 to 2.26 (when S10 is not considered due to its outlier behavior). All these features were also observed for point-to-point human hand movements (Engelbrecht (2001)). This conclusion shows on one side the validity of our measured data and on the other that we can make a first step towards a data-driven validation of SOC models for general goal-directed human movements by applying our ISOC algorithm.

Before results of this application are presented in the next section, the system equations (1) and (2) need to be set up for the driving experiment. The active steering wheel is modeled by the spring-damper dynamics

$$\Theta\ddot{\varphi} = -c\varphi - d\dot{\varphi} + M_{\mathrm{h}}, \tag{13}$$

where M_{h} is the steering torque applied by the human. It results from a second-order linear filter with time constants τ_1 and τ_2 describing the low-pass characteristics of the muscle dynamics (Winter (1990)):

$$\tau_1\dot{g} + g = u, \quad \tau_2\dot{M}_{\mathrm{h}} + M_{\mathrm{h}} = g, \tag{14}$$

where u is the neural activation and g the muscle excitation. We define the system state as $\boldsymbol{x} = [\varphi\ \dot{\varphi}\ M_{\mathrm{h}}\ g\ \varphi_{\mathrm{r}}]^\top$. The augmentation of it with the target steering angle φ_{r} (constant dynamics) is a mathematical reformulation to consider target steering angles $\varphi_{\mathrm{r}} \neq 0$ but still constitute the cost function in the form of (3):

$$J = \mathrm{E}\left\{ s_1(x_{N,1} - x_{N,5})^2 + s_2 x_{N,2}^2 + s_3 x_{N,3}^2 + \sum_{t=0}^{N-1} s_4 u_t^2 \right\}, \tag{15}$$

from which the definition of the cost function vector $\boldsymbol{s}$ follows. Discretizing (13) and (14) and choosing the numerical values as $\Theta = 0.056\,\mathrm{kgm}^2$, $c = 1.146\,\mathrm{Nm}$, $d = 0.859\,\mathrm{Nms}$ and $\tau_1 = \tau_2 = 0.04\,\mathrm{s}$ yields $\boldsymbol{A}$ and $\boldsymbol{B}$. Furthermore, $\boldsymbol{H} = [\boldsymbol{I}_{3\times3}\ \boldsymbol{0}_{3\times2}]$ is assumed for the human perception since φ_r only results from the above mentioned mathematical reformulation and g cannot be clearly associated with the human sensory apparatus. Due to lack of further knowledge about the additive noise processes and to keep the number of noise parameters tractable, independent stochastic processes $\alpha_{i,t}$ and $\beta_{i,t}$ are modeled for every state and output: $\boldsymbol{\Sigma}^{\boldsymbol{\xi}} = \mathrm{diag}([\sigma_1\ \sigma_2\ \sigma_3\ \sigma_4\ 0])$ and $\boldsymbol{\Sigma}^{\boldsymbol{\omega}} = \mathrm{diag}([\sigma_5\ \sigma_6\ \sigma_7])$. Finally, the signal-dependent noise processes are modeled by $\boldsymbol{C} = \sigma_8\boldsymbol{B}$ (single-input system), $\boldsymbol{D}_1 = \sigma_9\boldsymbol{H}\,\mathrm{diag}([1\ 0\ 0\ 0\ 0])$, $\boldsymbol{D}_2 = \sigma_{10}\boldsymbol{H}\,\mathrm{diag}([0\ 1\ 0\ 0\ 0])$ and $\boldsymbol{D}_3 = \sigma_{11}\boldsymbol{H}\,\mathrm{diag}([0\ 0\ 1\ 0\ 0])$ which completes the definition of the noise parameter vector $\boldsymbol{\sigma}$. Again, due to lack of further knowledge about the state-dependent noise processes, the $\boldsymbol{G}_i$ are chosen such that every perceived state is influenced by an independent noise process scaled by the corresponding state. In case of the LQG and LQ model, the signal-dependent and all noise parameters drop out, respectively.

For applying our ISOC algorithm, we first divide the measurement data into training and validation data. The parameters $\boldsymbol{s}$ and $\boldsymbol{\sigma}$ of the LQS, LQG and LQ model are identified for each subject individually with $\hat{\boldsymbol{m}}_t$ and $\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$ calculated from the repetitions of the $+\frac{2}{3}\pi$-step. Here, we set $\boldsymbol{M} = [\boldsymbol{I}_{2\times2}\ \boldsymbol{0}_{2\times3}]$ since φ and $\dot{\varphi}$ are measured. The repetitions of the $-\frac{2}{3}\pi$-step serve as validation data, i.e. the corresponding values for $\hat{\boldsymbol{m}}_t$ and $\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$ are the ground truth data for the LQS, LQG and LQ model predictions computed with $\tilde{\boldsymbol{s}}$ and $\tilde{\boldsymbol{\sigma}}$. The weighting vectors for the cost function bi-level optimization are chosen as $\boldsymbol{w}_\mathrm{m} = [0.9\ 0.9]^\top$ and $\boldsymbol{w}_\mathrm{v} = [0.1\ 0]^\top$ in case of the LQS and LQG model and as $\boldsymbol{w}_\mathrm{m} = [1\ 1]^\top$ and $\boldsymbol{w}_\mathrm{v} = [0\ 0]^\top$ for the LQ model. For the noise parameter bi-level optimization in case of the LQS and LQG model, $\boldsymbol{w}_\mathrm{m} = [0.1\ 0.1]^\top$ and $\boldsymbol{w}_\mathrm{v} = [0.9\ 0]^\top$ holds.

3. RESULTS

As motivated before, the focus of our model validation and comparison particularly lies on human average behavior. Hence, we first present detailed results of the LQS, LQG and LQ model in describing the human average steering behavior on training and validation data. Fig. 4 and Fig. 5 show the VAF values of all 14 subjects as well as the average over all subjects for $\mathrm{E}\{\varphi\}$ and $\mathrm{E}\{\dot{\varphi}\}$ on training and validation data, respectively. The LQS model outperforms the LQG and LQ model for every subject and in average on training and validation data, regarding $\mathrm{E}\{\varphi\}$ and $\mathrm{E}\{\dot{\varphi}\}$. On validation data, the LQS model predictions show VAF values of 0.989 ($\mathrm{E}\{\varphi\}$) and 0.893 ($\mathrm{E}\{\dot{\varphi}\}$) in average compared to 0.926 and 0.514 in case of the LQG and LQ model. When considering average behavior, the LQG and LQ model show identical predictions and VAF values as expected since the additive noise processes do not influence the predicted steering angle (velocity) mean (see Corollary 1). We perform an analysis of variance (ANOVA) to underline that the improvement of the LQS model is statistical significant by achieving p-values of $1.6\cdot10^{-4}$ and $2.5\cdot10^{-9}$ for $\mathrm{E}\{\varphi\}$ and $\mathrm{E}\{\dot{\varphi}\}$, respectively. On training data, very similar VAF and p-values result.

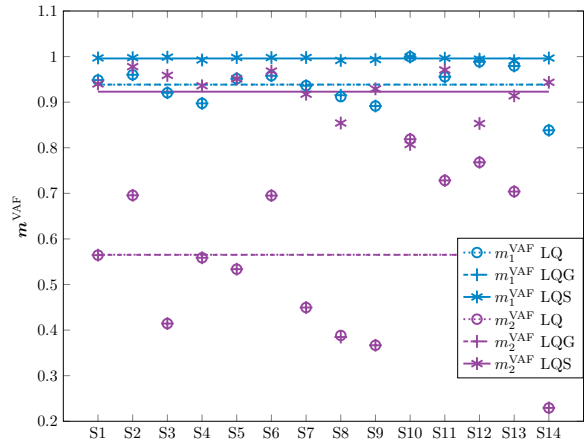

Fig. 4. VAF values for the steering angle mean (blue) and steering angle velocity mean (purple) achieved with the identified LQS, LQG and LQ model on the training data set. Lines correspond to the average value over all subjects.

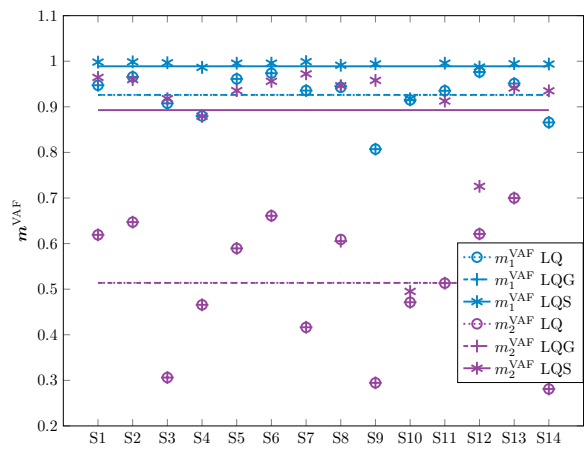

Fig. 5. VAF values for the steering angle mean (blue) and steering angle velocity mean (purple) achieved with the identified LQS, LQG and LQ model on the validation data set. Lines correspond to the average value over all subjects.

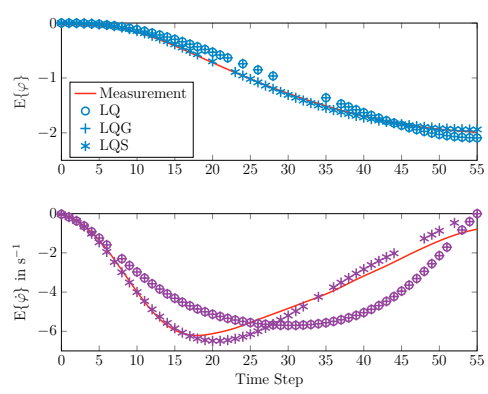

Fig. 6. Predictions of all three models compared to human measurement data on validation data for subject 2.

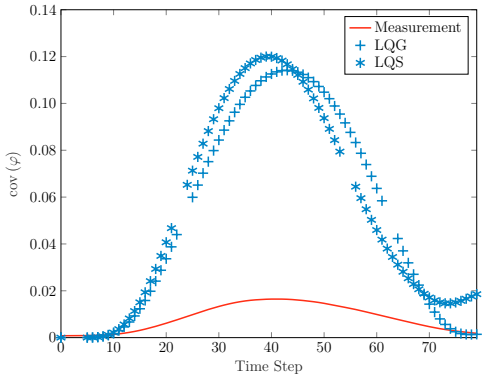

Fig. 7. Predictions of the LQG and LQS model compared to human measurement data on validation data for subject 1.

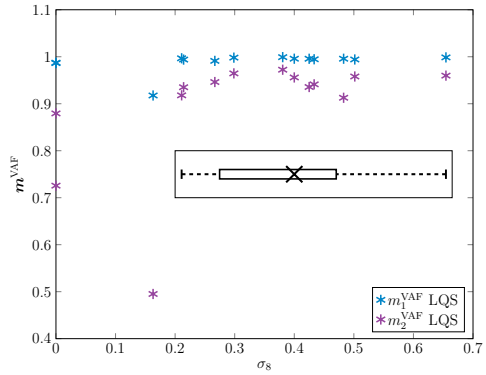

Fig. 8. VAF values achieved with the LQS model on validation data over control-dependent noise scaling parameter σ_8. Box plot for σ_8 is shown additionally where outliers at $\sigma_8 \approx 0$ and $\sigma_8 \approx 0.16$ are dropped.

In Fig. 6 the predictions of $\mathrm{E}\{\varphi\}$ and $\mathrm{E}\{\dot{\varphi}\}$ of the three models are depicted together with the human measurement for subject 2 to show the better performance of the LQS model qualitatively as well. Regarding the steering angle velocity, the LQS model especially describes the bell-shape and the maximum of the measurement curve better.

Concerning the steering angle variance, the human measurement data can only be described by the LQG and LQS model via overfitting. Both models achieve on the training data set VAF values from 0.451 to 0.996 with 0.880 in average (LQG) and from 0.616 to 0.985 with 0.904 in average (LQS), with no statistical difference. On validation data, the VAF values range from -104.2 to 0.895 with -7.288 in average (LQG) and from -107.1 to 0.894 with -7.648 in average (LQS). In Fig. 7, the predictions of the LQG and LQS model for subject 1 underline this poor performance in describing the measured steering angle variance qualitatively.

4. DISCUSSION

One of our main research objectives in this work is to analyze and prove the positive influence of the signal-dependent noise processes of the LQS model on describing human average behavior. Therefore, we begin the discussion of the results presented in Section 3 by formulating the following hypothesis.

Hypothesis 1. The LQS model describes human average steering behavior represented by the mean of steering angle $\mathrm{E}\{\varphi\}$ and steering angle velocity $\mathrm{E}\{\dot{\varphi}\}$ significantly better than a LQG or deterministic LQ model.

With Fig. 5 and the ANOVA results, Hypothesis 1 can be accepted since the LQS model outperforms the others for significance levels well below 1 %. Remarkably, the LQS model is able to describe the velocity profiles (see Fig. 3) better since the average VAF value for $\mathrm{E}\{\dot{\varphi}\}$ is nearly 75 % higher in case of the LQS model. Furthermore, the qualitative shape of the profiles matches better (see Fig. 6). However, $\mathrm{E}\{\dot{\varphi}\}$ is predicted slightly worse than $\mathrm{E}\{\varphi\}$ which can be concluded from the VAF values in Fig. 4 and Fig. 5 as well as from example subject 2 (see Fig. 6). The main reason for this is that the sections with constant reference φ_{r} were too short. Consequently, we set the threshold for determining the movement duration ($\dot{\varphi} < 0.1\,\mathrm{s}^{-1}$) relatively high to separate a single movement from a movement to a new target steering angle. Therefore, several velocity profiles do not show a convergence to zero at the movement ending which however the LQS model tends to predict (see Fig. 6). Furthermore, subject 10 seems to have problems in following these tough reference changes resulting in its outlier behavior (see Fig. 3) which none of the models can describe (VAF values < 0.5, see Fig. 5). Hence, with a reworked study design regarding the target steering angle we expect similar results for the predictions of $\mathrm{E}\{\varphi\}$ and $\mathrm{E}\{\dot{\varphi}\}$ in case of the LQS model as well as no outlier behavior of participants.

The significant better modeling performance of the LQS model compared to the reduced versions of it without signal-dependent noise processes is already a strong indication that especially these parts of the model have a positive influence on describing human average behavior. In order to provide a further evidence for this claim, Fig. 8 depicts an analysis of the scaling parameter of the control-dependent noise process which we assume to have the biggest positive influence according to findings in literature (see Harris and Wolpert (1998)). In Fig. 8, the VAF values for $\mathrm{E}\{\varphi\}$ and $\mathrm{E}\{\dot{\varphi}\}$ achieved on validation data with the identified LQS model are plotted over σ_8. Now, when the data points at $\sigma_8 \approx 0.16$ (subject 10) and at $\sigma_8 \approx 0$ (ISOC algorithm converged into a local optimum) are considered as outliers, for σ_8 the box plot in Fig. 8 results. The box plot highlights that the high VAF values (0.996 and 0.945 in average for $\mathrm{E}\{\varphi\}$ and $\mathrm{E}\{\dot{\varphi}\}$) of the remaining subjects are achieved with σ_8-values greater than 0.2. Hence, our data-driven analysis confirms findings in literature and theoretically expected by Corollary 1 that the signal-dependent and particularly the control-dependent noise process of the LQS model improves the performance of modeling human average behavior.

Concerning the description of human variability patterns with different SOC models, we cannot draw conclusions from our results. The main reason is that 14 repetitions of the same movement are (most likely) not enough to approximate the real human variability pattern. Hence, it cannot be assumed that the assumption on the estimate $\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$ in Problem 1 is fulfilled and the parameters determined on training data show poor generalization related

to covariance predictions. Fig. 7 yields a possible reason for this behavior. Although the measured variance of the steering angle in case of the shown $-\frac{2}{3}\pi$-step in the reference (validation data) shows reasonable characteristics, i.e. single-peaked curve with a maximum at the maximum velocity, the maximum value is with 0.0165 around one order of magnitude smaller than that on the training data (0.139). Since different but in the order of magnitude identical variance curves for the $+\frac{2}{3}\pi$-step and $-\frac{2}{3}\pi$-step would be expected, the estimated variance seems to be not representative for the real human variability. In future work, we will increase the number of single movements observed from subjects to overcome this problem. Since this poor variance description can already be seen at the steering angle, we dropped the velocity, which is only numerically computed from it, in the covariance fitting and analysis.

Although we look at a very specific application field (human steering behavior) at a first glance, the analysis of the kinematic features highlights that steering in our simplified driving task is indeed comparable to point-to-point human hand reaching (see Section 2.3). Hence, we strengthen on one side the proposition done by Kolekar et al. (2018) that steering is comparable to reaching and can make on the other a step towards a data-driven validation and analysis of SOC models for general goal-directed human movements. Thus, we are strongly convinced to be able to draw similar conclusions when we apply our ISOC algorithm to data of human hand movements: significant better performance of the LQS model in describing human average behavior quantitatively as well as qualitatively and scaling parameters of the control-dependent noise process as the main reason for this modeling performance.

5. CONCLUSION

In this paper, we use our previously developed Inverse Stochastic Optimal Control (ISOC) algorithm solving the identification problem of the linear-quadratic (LQ) sensorimotor (LQS) model to compare its capability in describing human steering behavior in a simplified driving task to a LQ Gaussian (LQG) and deterministic LQ model. In this context, an important focus lies on human average behavior. The final results show that the LQS model is able to predict the mean of steering angle and steering angle velocity with Variance Accounted For (VAF) values of 0.989 and 0.893 on a validation data set and averaged over 14 subjects. The improvement compared to the results of the LQG and LQ model is statistical significant with significance levels well below 1 %. Based on a parameter analysis, this significant better performance can be traced back to the control-dependent noise process characteristic for the LQS model. Since the movements in our simplified driving task show similar kinematic features as point-to-point human hand movements, this is a strong indication that the signal-dependent and in particular the control-dependent noise process of the LQS model is one of the crucial model parts to describe human average behavior. For human-machine systems, this better performance of the LQS model in characterizing human average behavior can be used for more accurate predictions of human movements in shared workspace applications or to facilitate a more intuitive support of humans in a trajectory tracking task.

REFERENCES

Engelbrecht, S. (2001). Minimum principles in motor control. *Journal of Mathematical Psychology*, 45, 497–542.

Flash, T. and Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5(7), 1688–1703.

Franklin, D.W. and Wolpert, D.M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, 72(3), 425–442.

Gallivan, J.P., Chapman, C.S., Wolpert, D.M., and Flanagan, J.R. (2018). Decision-making in sensorimotor control. *Nature Reviews. Neuroscience*, 19(9), 519–534.

Harris, C.M. and Wolpert, D.M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394, 780–784.

Karg, P., Kalb, L., Bengler, K., and Hohmann, S. (2021). Passing the baton between autopilot and driver. *ATZ Worldwide*, 123, 68–72.

Karg, P., Stoll, S., Rothfuß, S., and Hohmann, S. (2022). Inverse stochastic optimal control for linear-quadratic gaussian and linear-quadratic sensorimotor control models. *IEEE Conference on Decision and Control*, 2801–2808.

Kolekar, S., Mugge, W., and Abbink, D. (2018). Modeling intradriver steering variability based on sensorimotor control theories. *IEEE Transactions on Human-Machine Systems*, 48(3), 291–303.

MacAdam, C.C. (1981). Application of an optimal preview control for simulation of closed-loop automobile driving. *IEEE Transactions on Systems, Man, and Cybernetics*, 11(6), 393–399.

Markkula, G. (2014). Modeling driver control behavior in both routine and near-accident driving. *Human Factors and Ergonomics Society Annual Meeting*, 58(1), 879–883.

Nash, C.J. and Cole, D.J. (2015). Development of a novel model of driver-vehicle steering control incorporating sensory dynamics. *Dynamics of Vehicles on Roads and Tracks*, 72–81.

Ross, S. and Bagnell, J.A. (2010). Efficient reductions for imitation learning. *International Conference on Artificial Intelligence and Statistics*, 661–668.

Sentouh, C., Chevrel, P., Mars, F., and Claveau, F. (2009). A sensorimotor driver model for steering control. *IEEE International Conference on Systems, Man, and Cybernetics*, 2462–2467.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7(9), 907–915.

Todorov, E. (2005). Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Computation*, 17, 1084–1108.

Todorov, E. and Jordan, M.I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11), 1226–1235.

Uno, Y., Kawato, M., and Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics*, 61, 89–101.

Winter, D.A. (1990). *Biomechanics and Motor Control of Human Movement*. Wiley, 2nd edition.