

# Decision-making model for adaptive impedance control of teleoperation systems

Javier Corredor, Jorge Sofrony, Angelika Peer

**Abstract**—This paper presents a haptic assistance strategy for teleoperation that makes a task and situation-specific compromise between improving tracking performance or human-machine interaction in partially structured environments via the scheduling of the parameters of an admittance controller. The proposed assistance strategy builds on decision-making models and combines one of them with impedance control techniques that are standard in bilateral teleoperation systems. Even though several decision-making models have been proposed in cognitive science, their application to assisted teleoperation and assisted robotics has hardly been explored yet. Experimental data supports the Drift-Diffusion model as a suitable scheduling strategy for haptic shared control, in which the assistance mechanism can be adapted via the parameters of reward functions. Guidelines to tune the decision making model are presented. The influence of the reward structure on the realized haptic assistances is evaluated in a user study and results are compared to the no assistance and human assistance case.

**Index Terms**—Haptics, shared control, decision-making, assistance, telerobotics, drift-diffusion model



## 1 INTRODUCTION

A teleoperation system allows a human operator to perform a task remotely. Ideally this should permit the operator to perform the task as if s/he conducted it directly without any teleoperation system involved. Although human operators have the ability to abstract new information, learn and make decisions in unexpected situations, when they perform a task using teleoperation systems, their skills are not fully transferred to the remote site. This is mainly due to technical limitations in the telerobotic system, such as changes introduced in the visual feedback, delays in the signals exchanged between the operator and the remote site and restrictions of the robotic devices that complicate the interaction such as a restricted workspace or a limited number of degrees of freedom.

A way of reducing the effect of the aforementioned limitations is to provide assistance to the user performing the task. Typical challenges in this context are the selection of a proper assistance strategy and the finding of a suitable function that allows varying the level of the provided assistance between completely human-driven task execution and purely autonomous operation [1]. In this paper, we propose a haptic assistance mechanism that incorporates a decision-making model which varies the level of the provided assistance based on a compromise between improving tracking performance or human-machine interaction in partially structured environments.

The remainder of this paper is organized as follows: Section 2 provides a brief description of related work in the fields of haptic assistance and decision making models. This section is followed by the problem statement and an introduction to the newly proposed shared control approach based on decision-making models in Section 3. Next, Section 4 discusses the implementation of the haptic assistance. Section 5 presents the experimental setup and Section 6 summarizes results. Finally, in Section 7 conclusions are drawn on different parametrizations of the decision-making model and its relation to human and no assistance.

## 2 RELATED WORK

In robotics, autonomy can be defined in terms of a system's ability to function effectively without human intervention [2]. However for complex tasks, the system's implementation may be overly complicated and thus require task execution by humans. When such tasks are performed at distance, factors like fatigue may negatively affect the response and thus create the need for providing assistance.

The assistance strategy needs to be able to determine to what degree an autonomous agent should intervene to successfully execute the task. In human-robot interaction this mechanism is called assistance or arbitration, and allows to vary the authority or autonomy of task execution and is generally implemented as shared control strategy.

In this article we include models developed in the cognitive sciences field into a shared control scheme for teleoperation tasks in order to share the autonomy of task execution through the haptic channel. This section briefly presents relevant work in the haptic assistance and decision making model literature.

### 2.1 Haptic assistance

The haptic assistance term explicitly refers to the use of haptic cues to assist the human operator. These cues corre-

- J. Corredor and J. Sofrony are with the Universidad Nacional de Colombia, Department of System and Industrial Engineering and Mechatronics Engineering, Colombia.  
E-mail: [jcorredorc, jsufronye]@unal.edu.co
- A. Peer is with the Bristol Robotics Laboratory, University of the West of England, Bristol, UK and TUM-IAS Junior Fellow at the Institute of Advanced Studies of the Technische Universität München, Germany.  
E-mail: angelika.peer@brl.ac.uk

spond to forces generated artificially, which guide the user to successfully execute the task. These forces enable or prevent human movements in certain areas of the task space. Haptic assistance has been successfully implemented in many areas of physical human-machine interaction, including computer games [3], [4] or driver assistance systems [5]. Haptic assistance has been also implemented in mobile robots commanded at distance [6], [7], [8], [9], teleoperation interfaces [10], [11], haptic devices for training [12], [13], [14], and joint human-robot object transportation [15], [16], [17].

These studies commonly present human-machine interaction in structured environments, in which following a known path is usually the task at hand; this approach may be extended by including an estimation of the path [18], but in general the environment is assumed to remain unchanged. High-performance haptic interaction can be achieved by means of a constant assistance when the improvement of one objective is required [1]. Nevertheless, a constant assistance level (eg. [6], [12].) may not function well in partially structured environments because it does not adapt to unexpected situations. If a high assistance level is presented, the user may perceive a lack of control. On the other hand, if a low assistance level is presented, the user has more authority, but at the cost of less assistance.

Throughout this paper, the term *nominal task* is used to refer to the task description that can be modeled and thus a plan can be designed before its execution; the real task includes elements that can not be anticipated or modeled.

The trade-off between performance and control effort must be considered in the design of an optimal assistance mechanism in unstructured or partially structured environments [1], [5], [19], [20], [21]. In order to solve this trade-off, varying the assistance level is proposed as a mechanism to improve multiple design objectives. Reed et al. [22] have shown that a specialization of roles is assumed between humans in tasks with haptic feedback and that these roles change during task execution. Thus, incorporating such varying strategies in a human-robot interaction may lead to a more intuitive and human-like interaction [19].

The level of authority may be established by trial and error, or by means of pilot experiments with several subjects to determine the parameters of the assistance. For instance, the parameters may be determined by optimizing one design objective e.g. task performance [23], or by solving a multi-objective problem e.g. task performance, comfort and effort [1].

On-line adaptation of the assistance level has been proposed in [20] by heuristics implemented as state machines; in [4], [5], [17], [21] the construction of a scheduling function is reported. This function, commonly noted as  $\alpha$ , estimates the assistance level based on metrics of task performance [1], interaction with the user [9], [21] or any other task information [6], [14], [17]. Recently, concepts borrowed from game theory have been proposed in [15] to give the robot a more human-like reaction.

The benefits of having varying assistance schemes over constant schemes (or non-assisted configurations) are explored in [1], [5], [16], [24]. They conclude that incorporating assistance improves task performance compared to no assistance, but this is only true when the human operator agrees

with the proposed assistance strategy. On the other hand, varying assistance improves task performance compared to constant assistance, but the users adapt better to constant assistance [1], [16].

## 2.2 Decision-making Models (DMM)

Providing the robot with human-compatible representations and reasoning mechanisms, may help improve human-robot interaction as the human may find the robot behavior more predictable and thus, perceive the interaction as more natural. Although decision making in general is a very active research area, the usage of decision-making models to provide more natural human-robot interaction in shared control tasks is still in its early attempts.

So far, few concepts borrowed from game theory have been successfully applied to coordinate human and robot movements via the haptic channel. In [15] authors formulate a common cost function for task performance and an optimal control law for position and force based on *Nash's equilibria*. When human and robot disagree, the control law is adapted to minimize the error between the force exerted by the user and the optimal force, which allows the system to adapt to unexpected situations. Oguz et al. [3] proposed three negotiation models (also borrowed from game theory) considering concession, competition and tit-for-tat to solve collaborative but also conflictive situations in haptic interaction.

Haptic assistances must work under real conditions with high uncertainty, complexity, and in rapidly changing environments. As noted by Busemeyer et al. [25] a utility-based approach as adopted by methods based on Game Theory may become overly complex in such situations. An alternative approach consists of understanding the underlying decision mechanism, which may lead to a simpler and more coherent representation. Moreover, humans may diverge from utility functions, adopting non-optimal behavior.

In Cognitive Science strategies for decision making are studied as a sequence of interdependent decisions (decisions that are interrelated) [26], [27], [28] and human decision-making mechanisms are understood as making a choice when sufficient evidence is accumulated favoring one alternative over other possible choices [29]. Decision-making has been studied intensively using the so called *Two-Alternative Forced-choice Task* (TAFCT), which requires choosing between two options. Performed experiments show that humans adopt a sub-optimal strategy called the "*matching point*", at which both choices result in the same reward (payoff) (cf. (mp) in Fig. 1). Only few subjects find the optimal sequence that achieves better performance.

Several reward functions have been proposed to isolate and test specific behaviors such as exploring or exploiting resources in human decision-making experiments (Fig. 1 shows e.g. the matching shoulder structure) [26]. Please note that in these functions the matching point and the global optimum not necessarily coincide.

As a result of aforementioned experiments three different categories of models have been proposed to characterize human decision-making mechanisms. The first assumes that decisions are taken based on patterns observed over past rewards, see e.g. the WSL rule in [26] that evaluates the

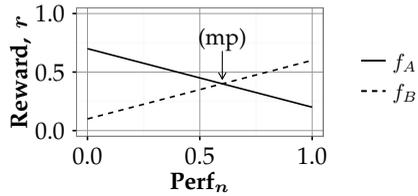


Fig. 1. Matching shoulders reward structure ( $f_A$ ,  $f_B$ ) [30]. In cognitive studies, the reward ( $r$ ) is calculated based on the amount of times an option was chosen in the last  $n$  decision making situations. Thus,  $\text{Perf}_n = \#A's/n$ , where "A" is the decision just made. If the human chooses option A, the reward is calculated based on  $f_A$ , otherwise, the reward is calculated based on  $f_B$ .

last two rewards received to take a decision on whether to keep the current selection or to change it. The second model assumes decisions to be taken based on the current reward and information such as the impact of the actual decision on future rewards and/or further attributes of choices. The most common attribute is the reward; but this approach may also include other attributes to make decisions e.g. expectations, beliefs or any kind of known information about alternatives. Thus, this kind of model can be seen as a model of the deliberation process<sup>1</sup>, see e.g. a Markov model approach in [32] or neural networks in [25]. The third model assumes that past rewards (evidence) are accumulated to determine the probability of future rewards for the various possible options and decisions to be taken on that basis to maximize the intake, an approach that has been shown to be used by humans when exploring unknown environments, see e.g. [33].

Among the models in the last category the DD model is popular because of its simplicity and success in fitting behavioral and neuronal data in several human studies (see [29] and references therein). The Drift-diffusion model fits the human behavior for experimental situations like interrogation and free response. In fact different models proposed in the cognitive sciences field adopt this behavior for parameters that reduce to the DD model [33].

The *Drift-Diffusion (DD) model* was included already in the supervision process of a team of robots by a human operator [30], [34] where the human changes his/her decisions in order to exploit or/and explore an environment. But, beside our own previous work [35] the DD model has not yet been explored in haptic shared control scenarios, which is the main purpose of this work.

### 3 PROBLEM STATEMENT AND APPROACH

Haptic assistance mechanisms help humans to improve task performance while guaranteeing a stable and transparent human-system interaction. Nevertheless, for high assistance levels, task performance is expected to be improved when the human agrees with the provided assistance. On the other hand, for low assistance levels little support is received making it easy to overrule the assistance and therefore task performance may decrease [5], [21].

1. Here deliberation is described as in [31] as the evolution of relative preference to options over time, then the decision rule can be thought of as a threshold level of preference required to stop deliberating and making a choice.

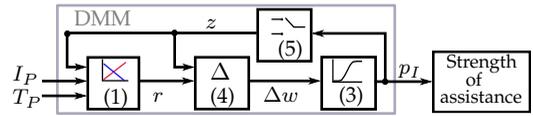


Fig. 2. The strength of the assistance is calculated based on the probability to improve some human-system performance metric, either  $T_P$  or  $I_P$ . The probability is a function of the accumulated evidence ( $\Delta$ ) of current and past rewards ( $r(t)$ ). Each block shows its corresponding equation number, namely: The reward structure (1), a reinforcement rule to accumulate the evidence (4), the drift-diffusion model (3) and the decision-making rule (5)

Haptic assistance design is a task-specific decision process, and may vary from task to task. We used a generic path-following task (the nominal task) with obstacles (unmodeled situations) in order to explore how the resulting haptic assistance varies when model parameters are changed. This approach allows understanding the importance of each parameter and its effects on human-robot interaction. This generic scenario underlies a series of real-life teleoperation applications (e.g. [18], [24]), but also other scenarios in physical human-robot interaction (consider e.g. a mobility assistant [36], human-robot collaborative manipulation [15], [16], rehabilitation [37]) or training with a haptic aid [38]. While in our considered scenario we assume the path to be known to the assistance, in reality this can be easily realized using a path planning algorithm and a map of the environment. The known obstacles represent obstacles that can be detected by the robot and that can consequently be considered in the path planning algorithm, while unknown obstacles represent situations in which the sensor system fails, e.g. due to changing lightning conditions.

The assistance strategy proposed builds on studies in the field of decision-making models and combines one of them with impedance control techniques that are standard in bilateral teleoperation systems. Having a minimum and a maximum strength for a haptic assistance is quite common and our proposed decision-making model helps taking a decision on the specific level of assistance provided by making a task-specific compromise between various performance measures. Trading-off between providing low or high assistance in haptic shared control tasks may namely be considered to follow similar decision-making mechanisms as used by humans in two-forced choice scenarios. In such scenarios the accumulated evidence for either choice is considered to be build on the actual and past rewards ( $r(t)$ ) with the evidence fluctuating until sufficient evidence is accumulated in favor of one of the alternatives. Since the probability of choosing a particular alternative depends on the accumulated evidence in favor of that alternative, we take advantage of this mechanism to vary the assistance level as function of this probability.

The purpose of this section is to review the components of the DMM negotiation strategy: the reward structure (1) and (2), the drift-diffusion model (3), the evidence accumulation (4) and the decision-making rule (5).

### 3.1 The reward structure

In cognitive experiments different reward structures are designed to provoke certain human behaviors in decision-making processes (e.g. exploit or explore). In our work, the reward structure is influenced to provide different assistance behaviors.

Although several reward functions have been proposed in cognitive science, a matching shoulder structure (see Fig. 1) has been chosen for our purposes to represent the known trade-off between task and interaction performance measures observed in the two situations with and without obstacles, representing modelled and unmodelled situations, see [1]. Although this trade-off may be also represented by non-linear functions as chosen by [36] and as observed e.g. by [1] for a task similar to the one studied in this manuscript, a linear approximation with contradicting slopes as in the original matching shoulder structure has been selected to simplify the design of the decision making mechanism built into the realized assistance function.

The reward functions are calculated based on human-system interaction metrics given by  $\text{Perf}_n \in [T_P I_P]$ , where  $T_P$  is a measurement of nominal task execution (i.e. nominal path tracking) and  $I_P$  is an interaction performance measurement which represents the additional physical effort experienced when human-system disagreement occurs. Assume that the reward  $r(t)$  is given by

$$r(t) = \begin{cases} r_I(I_P) & \text{if } z(t) = I, \\ r_T(T_P) & \text{if } z(t) = T, \end{cases} \quad (1)$$

where  $z(t)$  is the decision made at time  $t$ ,  $z(t) \in [T, I]$ ,  $T$  and  $I$  are the choices of improving nominal task performance or interaction performance, respectively. Variable  $r_T$  and  $r_I$  are the task and interaction reward functions (reward structure), respectively. The matching shoulder reward-type structure is given by

$$\begin{aligned} r_I &= k_I I_P + I_{P0}, \\ r_T &= k_T T_P + T_{P0}, \end{aligned} \quad (2)$$

where  $k_T$  and  $k_I$  are the ratios of change (slope) in the rewards over the change in the performances,  $T_{P0}$  is the highest task reward and  $I_{P0}$  is the lowest interaction reward. Functions  $r_T$  and  $r_I$  map the normalized performances to the rewards (for detail cf. Fig. 1 and Sec. 4.4.1). Generic examples for nominal task performance ( $T_P$ ) and interaction performance ( $I_P$ ) are considered in this paper, but can be customized when applying the approach to a specific application (for details see Sec. 4.3).

### 3.2 The decision-making model

The DD model makes decisions based on the accumulated evidence in favor of assisting the human to improve either of the proposed objectives, i.e. interaction or task performance. The soft-max model represents the DD behavior and is given by [39]

$$p_I(t+1) = \frac{1}{1 + \exp^{-\mu(w_I - w_T)}}, \quad (3)$$

where  $p_I$  is the probability to assist the human operator to improve human-system interaction in unmodeled situations,  $\mu$  is the slope of the sigmoidal function and  $w_I$

and  $w_T$  the evidence to improve interaction ( $I$ ) or task performance ( $T$ ), respectively. The evidence in favor of a specific performance measure (e.g. interaction) is noted as  $\Delta w = w_I - w_T$ .

The evidence in the DD-model-based assistance takes into account task and interaction performances. Therefore the (dynamic) rule for updating the evidence of the haptic assistance is given by

$$\begin{aligned} w_z(t+1) &= w_z(t) + \lambda(r_z(t) - w_z), \\ w_{\sim z}(t+1) &= \lambda r_{\sim z}(t), \end{aligned} \quad (4)$$

where  $\lambda \in [0, 1]$  is the learning rate and the symbol  $\sim$  denotes the "not" operator. Parameter  $\lambda$  represents the memory in the DD model: if  $\lambda$  is small, the model takes into account more history, this is, more memory; on the other hand, a larger  $\lambda$  means less memory, so the model takes into account only the most recent rewards to make the decision. The rule to update the weights (4) is inspired by the reinforcement rule equation in [39].

To evaluate the weight  $w_z$ , a choice has to be made based on some thresholding technique. The decision making rule is given by

$$z(t+1) = \begin{cases} I & \text{if } p_I > 0.5, \\ T & \text{otherwise.} \end{cases} \quad (5)$$

Note that  $\Delta w > 0$  indicates evidence in favor of  $I$ , while  $\Delta w < 0$  indicates evidence in favor of  $T$ . Since the decision follows the softmax rule (3) ( $p_I$ ), the threshold for the decision corresponds to  $p_I = 0.5$ .

The assistance level depends on the probability of improving one of the two performances (cf. (3)). Therefore, a scheduling function for the strength of the assistance  $\alpha \in [0, 1]$  may be given by

$$\alpha(p_I) = \frac{1}{2} + \frac{1}{2} \tanh\left(\frac{p_I - \phi}{\varphi}\right), \quad (6)$$

where,  $\phi$  is the switch point and  $\varphi$  is the smoothing level, both are user-defined parameters.

The  $\alpha$ -function maps the probability  $p_I$  to the assistance strength, noted as  $k$ . A linear homotopy is proposed to allocate the assistance levels,

$$k_{\perp} = \alpha k_{low} + (1 - \alpha) k_{high}. \quad (7)$$

The assistance varies the associated parameter to the assistance strength as a function of the probability ( $\alpha(p_I)$ ). Therefore, the assistance level can be adapted in a smooth and continuous fashion.

## 4 ONLINE ADAPTIVE ASSISTANCE

### 4.1 Control system architecture

The teleoperation setup and control system considered in this paper are standard (Fig. 3). A position-based admittance controller is used as the main mechanism to vary the assistance level perceived by the human operator (Fig. 4). It is assumed beforehand that the admittance-parameter variation is bounded and stable (cf. [35]). In order to present the proposed DD-model-based assistance behavior, a tracking task with unforeseen obstacles has been implemented. In the proposed task, the user has to track a desired path as

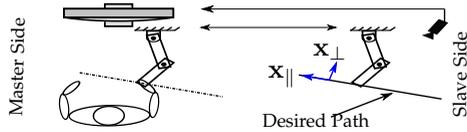


Fig. 3. Tracking task performed in the teleoperation scheme

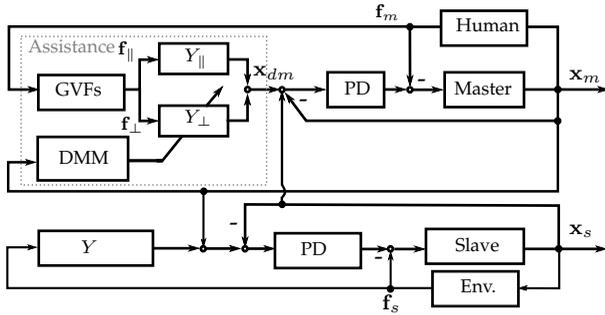


Fig. 4. Scheme of the decision-making assistance. The assistance consists of the decision-making model (DMM) and the guidance virtual fixtures (GVFs).

accurately as possible i.e. the nominal task (Fig. 3). The desired path is the only information from the environment that is available to the assistance; there is no available information about the location or dimensions of the obstacles. The operator receives additional visual information via video-link.

## 4.2 Implementation of the haptic assistance

The assistance is implemented as Guidance Virtual Fixtures (GVFs) [40] that force the system to stay on the desired path. The decision-making model varies the stiffness of the master admittance in the perpendicular direction of the desired path ( $Y_{\perp}$ ) as depicted in Fig. 4. Movements in the parallel direction were not constrained as the user should have full control over the task in this direction. A virtual fixture in this direction would drag or push the user, which would not be beneficial for task execution as task completion time plays no role in our task. Only some damping was added in order to prevent sudden movements.

The force measured at the tip of the master ( $\mathbf{f}_m$ ) can be decoupled as

$$\mathbf{f}_m = \mathbf{f}_{\parallel} + \mathbf{f}_{\perp}, \quad (8)$$

where  $\mathbf{f}$  is the force, indices  $\parallel$  and  $\perp$  represent the direction of application in the parallel and perpendicular direction of the desired path, respectively.

The admittance in the parallel direction ( $Y_{\parallel}$ ) and perpendicular direction ( $Y_{\perp}$ ) are given by

$$\mathbf{f}_{\parallel} = m_{\parallel} \ddot{\mathbf{x}}_{\parallel} + b_{\parallel} \dot{\mathbf{x}}_{\parallel}, \quad (9)$$

$$\mathbf{f}_{\perp} = m_{\perp} \ddot{\mathbf{x}}_{\perp} + b_{\perp} \dot{\mathbf{x}}_{\perp} + k_{\perp} \mathbf{x}_{\perp}, \quad (10)$$

where  $\mathbf{x}$  is the desired position,  $m$ ,  $b$  and  $k$  are the mass, damping and stiffness parameters, respectively. The desired position of the master is given by

$$\mathbf{x}_{dm} = \mathbf{x}_{\parallel} + \mathbf{x}_{\perp}. \quad (11)$$

## 4.3 Performance measures

Performance measures are calculated in the perpendicular direction of the path over an observation window of  $n = 100$  ms. In the normalization process, the maximum value for each observation window is updated during the interaction process. Nominal task performance is measured as the position error between the desired path and the current position of the master. The normalized version of the task performance measurement is given by,

$$T_{P,n} = \frac{\text{mean}_n(e)}{\max(e)},$$

where  $T_{P,n}$  is the normalized measurement and  $e$  is the tracking (position) error.

The interaction performance is defined by an agreement measure. The agreement is related to the internal forces between human and assistance, that when constrained to one degree of freedom, is given by

$$f_i = \begin{cases} f_{\perp} & \text{if } \text{sign}(f_{\perp}) \neq \text{sign}(f_{a\perp}) \wedge \|f_{\perp}\| \leq \|f_{a\perp}\|, \\ -f_{a\perp} & \text{if } \text{sign}(f_{\perp}) \neq \text{sign}(f_{a\perp}) \wedge \|f_{\perp}\| > \|f_{a\perp}\|, \\ 0 & \text{otherwise,} \end{cases} \quad (12)$$

where  $f_i$  is the internal force,  $f_{\perp}$  is the force exerted by the operator and  $f_a$  is the force exerted by the assistance. The normalized version of the internal forces ( $f_{i_n}$ ) is given by

$$f_{i_n} = \frac{\text{mean}_n(\|f_i\|)}{\max(\|f_i\|)}.$$

Because high internal forces mean low agreement, the agreement measurement is defined as

$$I_{P,n} = 1 - f_{i_n}. \quad (13)$$

Note that a similar result can be obtained by directly considering the internal forces ( $f_{i_n}$ ) as an interaction performance ( $I_{P,n}$ ) i.e. calculating  $I_{P,n}$  as function of disagreement ( $f_{i_n}$ ) instead of an agreement measure ( $1 - f_{i_n}$ ). For this, the reward  $r_I$  has to be redefined in (2), e.g.  $k_I = -0.5$  and  $I_{P0} = 1$ .

## 4.4 The Drift-Diffusion (DD) model parameters

Since an optimal choice of the involved parameters may be difficult to achieve, we will give some general guidelines on how to choose the parameters involved in the matching shoulder structure (i.e. slope and off-set of the reward functions) and the certainty-parameter  $\mu$  based on our selected generic example. Based on this generic example we explore how the resulting haptic assistance varies when changing model parameters giving the reader a better understanding of the importance of each parameter and its effects on results. When using the DD model for a specific application, the presented information can be used to customize an application-specific assistance.

Please note that the matching shoulder parameters,  $k_T, k_I, T_{P0}$  and  $I_{P0}$  in (2), regulate the rate of evidence accumulation, and the  $\mu$ -parameter in (3) configures the degree of certainty to improve a certain performance. The memory-parameter in the DD model (3) is assumed constant throughout ( $\lambda = 0.7$ ), hence reducing the number of parameters to be analyzed.

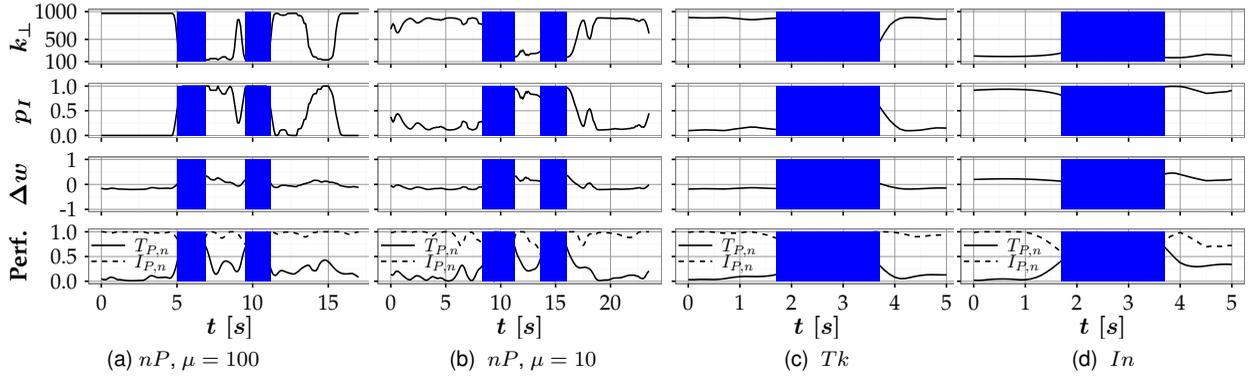


Fig. 5. Response to the Drift-Diffusion model-based assistance for a typical interaction with a human (see Subsection 5.1 for experimental setup). (a) For large  $\mu$  the user experiences a sudden change of stiffness, cf. (3). (b) For small  $\mu$  a smoother transition of stiffness; both (a) and (b) with values  $k_T = -1, k_I = 1, T_{P0} = 1$  and  $I_{P0} = 0$ . (c) The evidence ( $\Delta w$ ) to choose a stiffness that improves interaction performance ( $p_I$ ) is low, cf. (2) with values,  $k_T = -1, k_I = 0.5, T_{P0} = 1$  and  $I_{P0} = 0$ . (d) The evidence to assist in the interaction is high ( $\Delta w$ ), so the probability to assist the operator to improve the interaction is also high, cf. (2) with values  $k_T = -0.5, k_I = 1, T_{P0} = 0.5$  and  $I_{P0} = 0$ . Responses are recorded in the perpendicular direction of the desired path.

#### 4.4.1 Matching shoulder parameters

These parameters, cf. (2), influence the decision-making process of the haptic assistance. Note that equal slopes configure no preference ( $nP$ ) on any particular performance ( $k_I = k_T$ ) as both measures equally contribute to the accumulation of the reward. If the slope of the interaction reward function is reduced (e.g.  $k_I = 0.5, k_T = 1$  in Fig. 5c) the evidence in favor of interaction performance is accumulated at a lower rate compared to task performance; thus, a preference on task performance is configured. Likewise, to give weight to interaction performance, the slope for the reward of improving task performance is reduced (e.g.  $k_I = 1$  and  $k_T = 0.5$  in Fig. 5d).

When the task is preferred ( $Tk$ ) and the user moves on the path, the agreement and the tracking error accumulate evidence to assist in task performance. On the other hand, when the user moves off the path (cf. marked time window in Fig. 5c), the evidence to assist in improving interaction performance (agreement) increases slowly, as well as the probability  $p_I$ . When the user moves on the path s/he feels a higher level of assistance and when s/he moves off the path the stiffness decreases slowly, but generally the stiffness level remains quite high in favor of improving task performance.

If interaction is preferred ( $In$ ) and the user moves on the path, the agreement and the tracking error accumulate evidence to assist in improving interaction performance as evidence for improving task performance is accumulated at lower rate. When the user moves off the path (marked time window in Fig. 5d), the agreement decreases and the probability  $p_I$  is reduced because the evidence  $\Delta w$  is accumulated at a slow rate. The user feels a lower stiffness when s/he moves on the path while the stiffness slightly increases when the user moves off the path. The stiffness in average remains at low level in favor of improving interaction performance. Low assistance levels keep human-assistance agreement and the users easily adapt their movements to the proposed assistance.

#### 4.4.2 The $\mu$ parameter

In general, the  $\mu$ -parameter adapts the degree of certainty to improve a certain performance index. A large certainty

parameter ( $\mu$ -parameter) may configure the transition behavior as a switching mechanism between overassisted or underassisted in case the task demands such behavior. In the conducted experiment, a large certainty to improve interaction performance means that small movements outside the path are associated with a fast change in the levels of assistance (cf. marked time windows in Fig. 5a).

Such fast changes may be desirable when the operator requires to maintain control of the robot in the light of sudden changes in the remote environment; quick changes in the assistance level though may feel unnatural to the user, hence the user comfort may be reduced. On the other hand, if  $\mu$  is reduced to lower values (Fig. 5b), the operator does not experience sudden changes in the level of assistance, and the interaction may be perceived as more natural.

## 5 EXPERIMENT

We validate the DD-model-based assistance by studying the effects of different reward structure conditions on the assistance mechanism and by comparing their performance to human and no assistance. Our main hypotheses are:

- H1. DD-model-based assistance improves task performance compared to no assistance.
- H2. The DD-model-based assistance reproduces different assistance policies by means of the reward structure setup, resulting in:
  - Tracking error:  $Tk < nP < In$
  - Interaction performance:  $In < nP < Tk$
  - Physical dominance difference:  $In < nP < Tk$
  - Physical effort:  $In < nP < Tk$
 where  $Tk$ ,  $nP$  and  $In$  represent the reward setup for task preference, no preference and interaction preference, respectively.
- H3. One of the 3 configured variants of the DD-model-based assistance  $Tk$ ,  $nP$  and  $In$  setting different priorities for improving task or interaction performance should show more similarity with outcomes achieved by human assistance than others in terms of task performance, interaction performance, physical dominance and physical effort indicating that

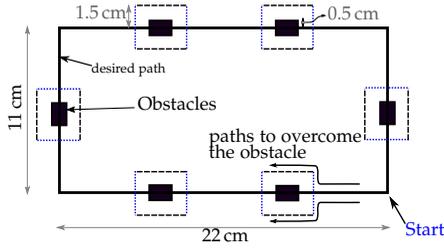


Fig. 6. Desired path indicated as solid line. A loop begins and ends at the point indicated with 'Start'. The suggested path to overcome the obstacle is shown.

humans may have adopted a similar policy for collaboration.

## 5.1 Apparatus and Experimental Setup

The assistance based on the decision-making model was implemented on a real teleoperation system with haptic feedback devices, see DeKiFeD and DeKiToP systems in [41]. The teleoperation system consists of an admittance-type haptic feedback device and teleoperator with four DOF; to validate the proposed assistance only three DOFs were used (z-direction is constrained and 3 planar joints are used to move device in horizontal plane keeping orientation of end-effector constant). The kinematics and technical specifications of the teleoperator are the same as the ones of the haptic input device, and a six DOF force/torque-sensor (JR3) was mounted at the tip of the manipulators to measure interaction forces with the human operator or the environment.

To determine a suitable virtual fixture admittance a task-dependent process is required [14]. The upper and lower bounds were determined in pilot experiments: a high level of assistance to reduce the nominal task error ( $k_{high} = 1000$  N/m) and a low level of assistance to reduce the disagreement when the user avoids the obstacles ( $k_{low} = 100$  N/m). These parameters represent lower and upper bounds because the DD-model adjusts the actual assistance strength based on  $p_I$  (3). The assistance is implemented using the following parameters,

$$\begin{aligned} m_{\parallel} &= 5 \text{ kg}, & b_{\parallel} &= 10 \text{ Ns/m}, \\ m_{\perp} &= 5 \text{ kg}, & b_{\perp} &= 200 \text{ Ns/m}, & k_{\perp} &= 100\text{-}1000 \text{ N/m}. \end{aligned}$$

The operator was asked to follow a desired path which was drawn on a piece of paper placed at the remote site and within the workspace of the slave (Fig. 6). Obstacles were added to the path to simulate unmodelled elements for the haptic assistance, while they are considered known to the human operator as a consequence of the visual feedback provided. The user was asked to avoid these obstacles following one of the two possible trajectories (passing the obstacle left or right, Fig. 6). No specific preference was induced for either of the two options. Participants were instructed to minimize the overall position error without paying attention to time. The introduced obstacles force the human to deviate from the original planned path (nominal task), which triggers the decision-making model of the haptic assistance as task performance and/or interaction performance suddenly change.

## 5.2 Measures

The effect of the DD-model-based assistance is evaluated by the Root-mean-square deviation (RMS) of the tracking error as task performance measure; the interaction is evaluated by the agreement of human and assistance and the physical dominance on task execution; as well as a power-based effort measure is considered to evaluate the effort of the dyad.

### 5.2.1 Tracking error

The root-mean-square of the tracking error ( $RMS(e)$ ) is considered as *task performance measure* for the interaction sequence. We analyzed movements on (nominal task) and off (unmodeled situations) the desired path separately. Doing so, a better understanding for the DD-model-based assistance behavior in these two situations can be gained.

The RMS is defined by,

$$RMS(e) = \sqrt{\frac{\sum_{k=1}^N (x_{n,k} - x_{m,k})^2}{N}}, \quad (14)$$

where  $k$  is the sample ( $k = [1, \dots, N]$ ),  $x_{n,k}$  is the desired position defined in the nominal task description,  $x_{m,k}$  the actual position of the master device and  $N$  the number of samples in the examined interaction sequence.

### 5.2.2 Internal forces

Internal forces were calculated as already defined in (12), Sec. 4.3. The mean of internal forces describes the disagreement between human and assistance. Again, we analysed movements on and off the path separately.

### 5.2.3 Physical Dominance Difference

The partner in a dyad who applies higher manipulation force commands the object movement to a higher degree and can thus be considered dominant [19].

Internal forces ( $f_i$ ) occur when the two partners do not apply forces in the same direction. These forces are related to no-work forces [19]. On the other hand, the physical dominance of partner 1 over 2 ( $PD_{12}$ ) is defined by the external force which generates work as follows [19],

$$PD_{12,k} = \frac{f_{e1,k}}{f_{sum,k}},$$

where  $f_{sum,k}$  is the sum of external forces and  $f_{e1,k}$  is the external force exerted by partner 1, which can be calculated by

$$f_{e1,k} = f_{1,k} - f_{i1,k}, \quad (15)$$

$$\begin{aligned} f_{sum,k} &= f_{e1,k} + f_{e2,k}, \\ &= f_{1,k} + f_{2,k}, \end{aligned} \quad (16)$$

where  $f_{1,k}$  is the force exerted by partner 1,  $f_{i1,k}$  the internal force, and (16) implies that  $f_{i1,k} \equiv -f_{i2,k}$ .

A partner is absolutely dominant with a value of one, and absolutely non-dominant with a value of zero and  $PD_{12} \in [0, 1]$  and  $PD_{12} + PD_{21} = 1$ . The amount of dominance difference ( $PD_{diff}$ ) means the amount to which one partner dominates the other and is derived by [19],

$$PD_{diff} = |\overline{PD}_{12} - \overline{PD}_{21}|,$$

where  $\bar{\cdot}$  is the symbol of the mean. A value of zero of  $PD_{diff}$  means that no-dominant partner exists in the interaction. On the other hand, a non-zero value means that a dominant partner exists.

Note that internal forces represent an interaction measure in which no movement of the object is produced. Besides, physical dominance involves the external forces that generate movement and thus, determine the control of the jointly handled object. Both measures complement the analysis in a joint dyadic haptic interaction task.

#### 5.2.4 Power-based effort

Higher energy flows correlate with a higher physical effort. Thus, the effort measure is calculated as a function of the power  $P_*$  exchanged

$$P_* = \dot{\mathbf{x}} \mathbf{f}_*$$

where  $P_*$  is the power from partner ‘\*’ to the environment (including the assistance),  $* \in [1, 2]$ ,  $\dot{\mathbf{x}}$  the velocity of the object and  $\mathbf{f}_*$  is the force applied by the partner ‘\*’. The velocity is equivalent for both partners when they hold on to the same interaction point.

A positive energy flow, i.e. energy injection to the system (e.g. by accelerating the virtual object) causes physical effort for the operator, but also a negative energy flow results in physical effort as energy needs to be dissipated (e.g. by decelerating the virtual object).

Our effort measure captures the total effort of human and assistance and can be calculated by [19],

$$MAP_T = MAP_1 + MAP_2 = \frac{1}{N} \sum_{k=1}^N |P_{1,k}| + \frac{1}{N} \sum_{k=1}^N |P_{2,k}| \quad (17)$$

where,  $P_{1,k}$  and  $P_{2,k}$  are the energy flows at the respective interface points at a given time step  $k$ .

### 5.3 Conditions

As the principal goal is the evaluation of the DD model in the haptic assistance scenario, three different conditions which explore the influence of the matching shoulder reward structure parameters in the haptic assistance were considered: **1)  $nP$** , the decision-making model has no preference for any performance measure,  $k_T = 1, k_I = 1, T_{P0} = 1, I_{P0} = 0$ ; **2)  $Tk$** , task performance is favored,  $k_T = -1, k_I = 0.5, T_{P0} = 1, I_{P0} = 0$ ; and **3)  $In$** , interaction performance is favored,  $k_T = -0.5, k_I = 1, T_{P0} = 0.5, I_{P0} = 0$ . In order to compare with human-like execution schemes, we further introduced a condition **4)  $Ha$** , with an expert operator haptically coupled with the tested subject, cf. e.g. [42] and a condition **5)  $Na$** , with one operator performing the task alone and without any assistance. Although human-human interaction in a teleoperation task may produce larger position errors compared to natural human-human interaction, we propose such a comparison as it presents a possible reference for designing more natural interaction strategies [19].

The expertise level of the users in interacting with a haptic assistance while carrying out this particular task may change the outcome of the experiment as it may affect the performance level and ability to adapt to a provided

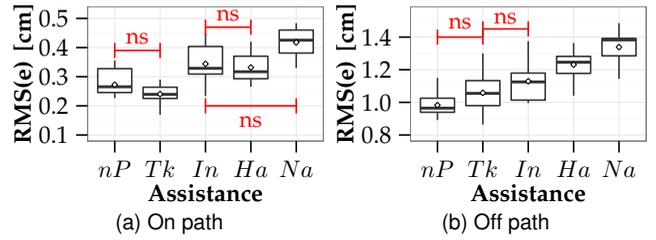


Fig. 7. Boxplot for the RMS of tracking error (samples divided for on and off path). No significant differences between groups are marked by ‘ns’ meaning  $p > 0.05$ . All other non-marked comparisons are significant on a  $p = 0.05$  level.

assistance. Hence, in order to reduce this bias unexperienced test subjects not having performed the particular assisted task before were chosen. An expert user though having practiced the task beforehand in interaction with other training persons not participating in the experiment takes the role of the haptic assistance in the human-human condition.

### 5.4 Participants and procedure

Twelve healthy subjects were tested and asked to perform the same task under the five aforementioned conditions. Two operators consider themselves to have an excellent experience in the use of teleoperation systems, four operators a good experience, four operators a regular experience and two operators admitted no experience. However none of the operators had previous experience with the interface.

At the beginning of the experiment, the operator executed a trial run in order to get familiar with the teleoperation interface. Then, five different assistance strategies were presented to the operator in random order to reduce learning effects. The operator was unaware of the current conditions presented, except when the  $Ha$  assistance was selected. In the  $Ha$  assistance, the operator was informed that the expert partner has the same degree of expertise to avoid taking him/her a preference on a passive role in the execution of the task.

## 6 RESULTS AND DISCUSSION

Since performance measures were not normally distributed (according to the Kolmogorov Smirnov test), a non-parametric *one factorial repeated-measurements analysis* of variance (ANOVA) was performed (namely the Friedman test), and significant differences were evaluated at  $p = 0.05$  level. In addition, all *post-hoc* tests were conducted using a Wilcoxon test with Holm-Bonferroni correction; the p-values for these tests are presented in Tables 1, 2, 3 and 4. For selected comparisons we also performed a two one-sided test for similarity<sup>2</sup> with Holm-Bonferroni correction.

**Tracking error:** When the user moves on the path (Fig. 7a and Table 1), the DD-model-based assistance conditions improve task performance compared to  $Na$  ( $p < 0.05$ ), except for  $In$  ( $p = 0.081$ ).

Further, the DD-model-based assistance improves nominal task performance compared to  $Ha$  ( $p < 0.05$ ); except for  $In$  ( $p = 0.62$ ). A similarity test between  $Ha$  and  $In$  turned out to be significant ( $p < 0.05, \epsilon = 0.01$ ).

2. see R-package ‘equivalence’ command `rtost`

TABLE 1  
p-value ( $p$ ) and effect size ( $r$ ) for RMS error when human moves on and off the path

		$nP$		$Tk$		$In$		$Ha$	
		$p$	$r$	$p$	$r$	$p$	$r$	$p$	$r$
On	$Tk$	0.185	0.330	-	-	-	-	-	-
	$In$	<b>0.007</b>	0.436	<b>0.005</b>	<b>0.731</b>	-	-	-	-
	$Ha$	<b>0.049</b>	0.495	<b>0.005</b>	<b>0.801</b>	0.622	0.130	-	-
	$Na$	<b>0.015</b>	<b>0.778</b>	<b>0.005</b>	<b>0.849</b>	0.081	<b>0.530</b>	<b>0.017</b>	<b>0.660</b>
Off	$Tk$	0.068	0.306	-	-	-	-	-	-
	$In$	<b>0.005</b>	<b>0.625</b>	0.129	0.259	-	-	-	-
	$Ha$	<b>0.005</b>	<b>0.813</b>	<b>0.005</b>	<b>0.601</b>	0.037	0.460	-	-
	$Na$	<b>0.005</b>	<b>0.837</b>	<b>0.006</b>	<b>0.778</b>	<b>0.017</b>	<b>0.707</b>	<b>0.037</b>	<b>0.554</b>

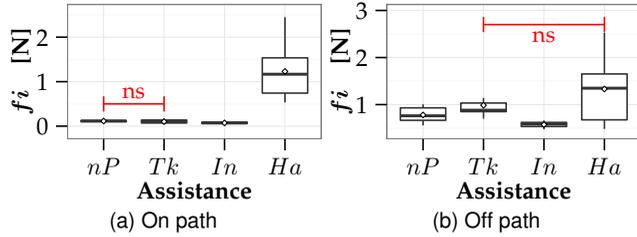


Fig. 8. Boxplot for the internal forces. No significant differences between groups are marked by 'ns' meaning  $p > 0.05$ . All other non-marked comparisons are significant on a  $p = 0.05$  level.

Comparing the DD-model-based assistances, we observed that  $nP$  and  $Tk$  show the best task performance behavior when the user moves on the path (Fig. 7a). In addition,  $In$  assistance achieves the lowest task performance among all DD-model-based-assistances for "on path".

When the user moves off the path we found that the position error is larger for  $Na$  compared to all other conditions, followed by  $Ha$  (Table 1 and Fig. 7b). For  $nP$  and  $Tk$  no significant differences with respect to task performance were found when the user moves off the path ( $p = 0.068$ ); neither between  $In$  and  $Tk$  ( $p = 0.129$ ).

**Internal forces:** In both movements, on and off the path, the DD-model-based assistance reduces internal forces compared to  $Ha$  (Table 2 and Fig. 8).

Regarding the on-path movements, no significant difference was found in terms of agreement between  $nP$  and  $Tk$ . Furthermore, statistically the lowest internal forces were presented for  $In$  assistance (Table 2 and Fig. 8a). When the user moved off the path, internal forces kept low values for DD-model-based assistance. The results show that the lowest internal forces were obtained for  $In$ , followed by  $nP$  and  $Tk$  (Table 2 and Fig. 8b).

TABLE 2  
p-value ( $p$ ) and effect size ( $r$ ) for internal forces ( $f_i$ ) when the user moves on and off the path

		$nP$		$Tk$		$In$	
		$p$	$r$	$p$	$r$	$p$	$r$
On	$Tk$	0.424	0.153	-	-	-	-
	$In$	<b>0.003</b>	<b>0.742</b>	<b>0.024</b>	0.365	-	-
	$Ha$	<b>0.003</b>	<b>0.849</b>	<b>0.003</b>	<b>0.849</b>	<b>0.003</b>	<b>0.849</b>
Off	$Tk$	<b>0.037</b>	0.354	-	-	-	-
	$In$	<b>0.034</b>	<b>0.625</b>	<b>0.003</b>	<b>0.849</b>	-	-
	$Ha$	<b>0.037</b>	<b>0.377</b>	0.092	0.200	<b>0.034</b>	<b>0.636</b>

**Physical Dominance Difference:** Statistical evidence on physical dominance difference measures reveals differences between  $Ha$  and  $nP$  ( $p = 0.003$ ), between  $Ha$  and  $In$  ( $p = 0.003$ ), as well as between  $Tk$  and  $In$  ( $p = 0.003$ ).

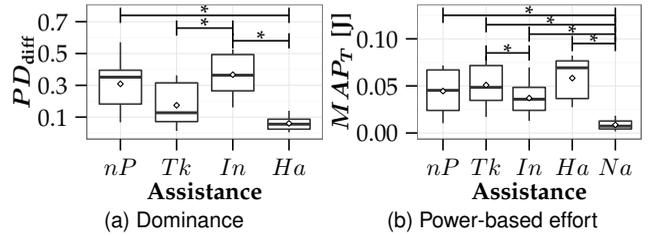


Fig. 9. Boxplot for the dyadic power-based effort and physical dominance difference. Significant differences between groups are shown with a (\*) corresponding to  $p = 0.05$  level of significance.

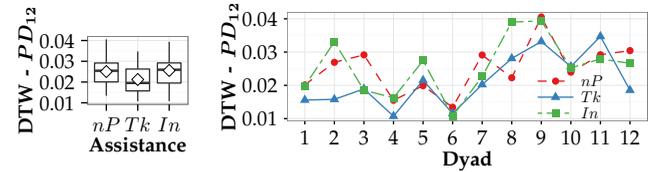


Fig. 10. Boxplot and values for each dyad for the dynamic time warping (DTW) distance between the  $Ha$  and DD-model-based assistances for physical dominance of human over assistance ( $PD_{12}$ ).

Furthermore, no-differences were found between  $Ha$  and  $Tk$  ( $p = 0.081$ ), between  $nP$  and  $Tk$ , nor  $nP$  and  $In$  (Fig. 9a and Table 3).

TABLE 3  
p-value ( $p$ ) and effect size ( $r$ ) for physical dominance difference ( $PD_{diff}$ )

	$nP$		$Tk$		$In$	
	$p$	$r$	$p$	$r$	$p$	$r$
$Tk$	0.081	0.448	-	-	-	-
$In$	0.176	0.153	<b>0.003</b>	<b>0.577</b>	-	-
$Ha$	<b>0.003</b>	<b>0.754</b>	0.081	0.460	<b>0.003</b>	<b>0.849</b>

**Power-based effort:** In terms of dyadic effort  $In$  exhibited a reduction compared to  $Tk$  ( $p = 0.021$ ). On the other hand, no differences were found between  $nP$  and  $In$ , nor between  $nP$  and  $Tk$  (Table 4 and Fig. 9b). Comparisons between DD-model-based assistances and  $Ha$  showed no significance, but a similarity test between  $Ha$  and  $Tk$  resulted to be significant ( $p = 0.005$ ,  $\epsilon = 0.05$ ).

We further included the condition  $Na$ , an individual case, with the purpose of comparing the effort exerted by the user with no assistance and with the proposed assistances (Fig. 9b). The  $Na$  condition presents the lowest effort, but also the lowest task performance compared with the dyadic assistances ( $p < 0.01$ ).

TABLE 4  
p-value ( $p$ ) and effect size ( $r$ ) for power-based effort ( $MAP_T$ )

	$nP$		$Tk$		$In$		$Ha$	
	$p$	$r$	$p$	$r$	$p$	$r$	$p$	$r$
$Tk$	0.170	0.165	-	-	-	-	-	-
$In$	0.467	0.141	<b>0.021</b>	0.318	-	-	-	-
$Ha$	0.388	0.424	0.467	0.259	0.081	0.495	-	-
$Na$	<b>0.005</b>	<b>0.766</b>	<b>0.005</b>	<b>0.837</b>	<b>0.005</b>	<b>0.790</b>	<b>0.005</b>	<b>0.849</b>

**Time series analysis:** Since a times series analysis may provide deeper insight into the similarity of DD-model-based assistances and the human assistance condition ( $Ha$ ), we implemented the Dynamic Time Warping (DTW) algorithm [43] to compare the physical dominance of the user

over the assistance ( $PD_{12}$ ). We evaluated the similarity of the  $Ha$  condition with the three conditions  $nP$ ,  $Tk$ , and  $In$ , see Figure 10 for results.

A one way repeated measures ANOVA with Holm-Bonferroni correction, showed no significant differences between the DTW-distance measures comparing the  $Ha$  and the DD-model-based assistance schemes. Figure 10 further illustrates the DTW-distance measures obtained for individual dyads and for the three conditions  $nP$ ,  $Tk$ , and  $In$ .

## 6.1 Discussion

### H1. DD-model-based assistance improves task performance compared to no assistance

In the on-path movements the low assistance level in  $In$  cannot provide the required assistance level to improve nominal task performance compared to  $Na$ . In contrast,  $nP$  and  $Tk$  provide a high assistance level in the nominal task and thus, task performance is improved compared to  $Na$ .

In off-path situations all DD-model-based assistances improve task performance with respect to  $Na$ , but at different rates with  $nP$  performing best. Further, in off-path situations the  $RMS$  value of the error across all conditions is greater than the dimensions of the obstacle<sup>3</sup>, this suggests that users, on average, avoid successfully the obstacles in all conditions.

Thus, H1 can be partially confirmed. Despite  $nP$  and  $Tk$  improve task performance compared to  $Na$  in on-path situations,  $In$  fails to show any improvement compared to  $Na$ . In off-path situations all DD-model-based assistances improve task performance with respect to  $Na$ .

In telerobotics, Boessenkool et.al [24] reported that haptic assistance improves task performance compared to no-assistance in a structured environment; in our case a similar result is obtained but extended to a partially structured environment.

### H2. DD-model-based assistance reproduces different assistance policies depending on reward structure setup

Regarding task performance in on-path movements,  $Tk$  and  $nP$  set the best task performance, while for  $In$  the configuration of the reward function provides low assistance levels allowing the user to adopt free movements.

When the human moves off the path using either  $In$  or  $Tk$  assistance strategies, the control authority is handed to the user but at different rates. For instance,  $In$  provides low assistance as the reward structure has been designed to favor interaction performance, hence the user can freely choose with what distance he/she passes obstacles. In the  $Tk$  condition, higher assistance levels were provided, and therefore the human had to fight against the assistance while circumventing obstacles, leading to the effect that users tried to keep some safety distance to obstacles. Only, when both criteria were balanced as in the  $nP$  condition, users passed obstacles very closely.

Results further show that  $Tk$  produced higher internal forces compared to  $nP$  (Table 2 and Fig. 8b), irrespective of

whether the user moved on or off the path. This suggests that task performance is improved, while the agreement is decreased in the  $Tk$  condition, while exactly the reverse could be observed for the  $In$  condition.

The high assistance level in  $Tk$  causes that the user interacts actively with the proposed assistance leading to a more equal share of the workload. On the contrary, low assistance levels in the  $In$  condition, give task control to the user, see (Fig. 9a and Table 3), while  $nP$  can be found inbetween. This suggests that the reward functions successfully allocated the physical dominance in the DD-model-based assistances.

Regarding physical effort, condition  $In$  with its low assistance level leads to the least effort among the three tested DD-model-based assistances, because the user does not have to *fight* against the assistance. The  $Tk$  condition leads to the highest effort and the  $nP$  condition can be found somewhere inbetween.

Thus, overall we can confirm H2 and with it the influence of the reward structure on the haptic assistance as their effect on task performance, interaction performance, physical dominance and physical effort is as expected.

### H3. One of the DD-model-based assistances shows higher similarity to human assistance than others

Regarding task performance (in on-path movements)  $In$  performs similar to  $Ha$  (Fig. 7a), while  $nP$  and  $Tk$  significantly improve this metric with respect to  $Ha$ . Regarding off-path movements all DD-model-based assistances perform better in terms of task performance than  $Ha$ .

While all DD-model-based assistances successfully reduced internal forces compared to  $Ha$  in the on-path condition (Fig. 8b), for  $Tk$  no significant difference compared to  $Ha$  could be found for off-path conditions.

In general, physical dominance is shared between human and assistance in all conditions – the  $PD_{diff}$  value is less than 0.5. Nevertheless,  $Ha$  is clearly the condition with the most equal physical dominance strategy, and it is naturally adopted by the users. This can be explained by the collaborative strategy adopted by the collaborator, and may also suggest that the user participates actively in task execution, either to collaborate or fight against the collaborator. Furthermore, high assistance levels in the  $Tk$  condition also let on average to a rather equal distribution of dominance. This might suggest that dominance in the  $Tk$  assistance is shared between user and assistance in the sense that, the human avoids obstacles while  $Tk$  tries to maintain the movements on the path. It is also interesting to note that no statistical difference was found between the  $Ha$  condition and any of the DD-model-based assistances in terms of physical effort, while similar performance was found between  $Ha$  and  $Tk$ .

Overall, H3 cannot be confirmed based on results representing means over all subjects. Significant differences between  $Ha$  and DD-model-based assistances were found for task performance, interaction performance, and dominance difference. The  $Tk$  condition performed closest to the  $Ha$  condition in terms of interaction performance and physical dominance difference, while the  $In$  condition performed closest for task performance. Considering the analysis based on means collaborating humans in the  $Ha$  condition achieved an overall more balanced physical domi-

3. The perpendicular dimension of the obstacle is 0.5 cm. On average, the error when the user moves off path for each condition is greater than 0.9 cm, cf. Fig. 6

nance, used higher interaction forces and achieved a worse task performance compared to the realized DD-model-based assistances. Only in terms of effort  $Ha$  and  $Tk$  showed similarities. Thus summarizing, on average the realized DD-model-based assistances seem to be more compliant than the average human collaborator as they leave more physical dominance with the user and thus, allow achieving an overall better task performance. Efforts spend in the  $Ha$  condition on fighting with the collaborator, seem to be invested in the DD-model-based assistance into the improvement of task performance. This provides evidence that a slightly unequal dominance distribution as realized in the DD-model-based assistances may benefit task performance.

A more detailed analysis involving a time-series analysis on individual dyad level and involving similarity measures by applying Dynamic Time Warping of observed outcomes, provides a deeper insight and shows slightly different results. While the mean over all DTW-distance measures does again not provide further insight, looking at the similarity measures of individual dyads indicates that they may have adopted different policies resulting in different dominance patterns (most of them adopting a policy that resulted in dominance measures most similar to the  $Tk$  condition) making an interpretation based on means difficult. This, highlights the need for a time-series analysis on individual level as analyzing means over dyads can hide effects if they behave differently.

## 7 CONCLUSION

We have proposed a novel haptic assistance that enhances tracking performance and human-machine interaction in a teleoperation setup via the scheduling of admittance control parameters when unmodeled situations occur in the nominal task. The *Drift-Diffusion Model* was included in the haptic assistance as the decision-maker of the specific sharing strategy. Our results show that the Drift-Diffusion model is a suitable scheduling strategy for haptic shared control or other control allocation tasks as demonstrated by experimental data. All implementations were found to improve performance over the no assistance case. Furthermore, the parameters of the reward structure were found to influence the decision making process, which in turn adapts the assistance for different task requirements. However, significant differences between the performance of the Drift-Diffusion-model-based assistances and the mean performance of dyads was observed that were found to operate at rather equal physical dominance levels, indicating that the dominance distribution plays an important role and an unequal distribution may benefit overall task performance. The large variance in measures, however, also indicated that a more detailed analysis on individual dyad level and involving similarity measures was required to provide a deeper insight. Analyzing the dominance distribution, dyads were found to adopt different policies, most of them applying a policy resulting in outcomes showing most similarity with the task-performance optimizing implementation of the DD-model-based assistance.

Although path following and obstacle avoidance may be considered generic tasks, the proposed decision making algorithm has to be adjusted to the particular application

by properly selecting reward functions and their parameters. While in this manuscript we have provided general guidelines on how to decide on reward functions and their parametrization, more formal algorithms for either extracting adopted policies and their parametrization from observation of human-human collaboration or their optimization based on optimization routines or reinforcement learning approaches may be investigated in future work.

Further, future work may focus on adapting the level of assistance based on an online adaptation of the slope of the matching shoulder structure depending on new incoming information on the task or user, the scheduling between more than two levels by including voting mechanisms as well as the exploration of alternative decision-making models that e.g. include decision making processes of the partner into the assistance mechanism.

## ACKNOWLEDGMENT

This work is supported in part by the EU project MOBOT within the 7th Framework Programme of the European Union, contract number ICT-2013-600796 as well as the Institute of Advanced Studies of the Technische Universität München. The doctoral studies of the first author and his visit to Technische Universität München were supported by the National PhD Support Program, Instituto Colombiano para el desarrollo de la Ciencia y la Tecnología (COLCIEN-CIAS), Colombia.

## REFERENCES

- [1] C. Passenberg, A. Glaser, and A. Peer, "Exploring the design space of haptic assistants: the assistance policy module," *IEEE Transactions on Haptics*, vol. Early Access Online, 2013.
- [2] M. B. Dias, B. Kannan, B. Browning, E. G. Jones, B. Argall, M. F. Dias, M. Zinck, M. M. Veloso, and A. J. Stentz, "Sliding autonomy for peer-to-peer human-robot teams 1," in *10th International Conference on Intelligent Autonomous Systems (IAS 2008)*, 2008.
- [3] S. Oguz, A. Kucukyilmaz, T. Sezgin, and C. Basdogan, "Supporting negotiation behavior with haptics-enabled human-computer interfaces," *IEEE Transactions on Haptics*, vol. 5, no. 3, pp. 274–284, 2012.
- [4] A. Enes and W. Book, "Blended Shared Control of Zermelo's navigation problem," in *American Control Conference (ACC)*, 2010, Jun. 2010, pp. 4307–4312.
- [5] D. A. Abbink, M. Mulder, and E. R. Boer, "Haptic shared control: smoothly shifting control authority?" *Cogn. Technol. Work*, vol. 14, no. 1, p. 1928, Mar. 2012.
- [6] M. Desai and H. Yanco, "Blending human and robot inputs for sliding scale autonomy," in *IEEE International Workshop on Robot and Human Interactive Communication*, 2005. ROMAN 2005, Aug. 2005, pp. 537–542.
- [7] D. Bruemmer, D. Few, R. Boring, J. Marble, M. Walton, and C. Nielsen, "Shared understanding for collaborative control," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 35, no. 4, pp. 494–504, Jul. 2005.
- [8] J. Shen, J. Ibanez-Guzman, T. C. Ng, and B. S. Chew, "A collaborative-shared control system with safe obstacle avoidance capability," in *2004 IEEE Conference on Robotics, Automation and Mechatronics*, vol. 1, Dec. 2004, pp. 119–123 vol.1.
- [9] J. Medina, D. Lee, and S. Hirche, "Risk-sensitive optimal feedback control for haptic assistance," in *IEEE International Conference on Robotics and Automation*, 2012, pp. 1025–1031.
- [10] C. Passenberg, A. Peer, and M. Buss, "A survey of environment-, operator-, and task-adapted controllers for teleoperation systems," *Mechatronics*, vol. 20, no. 7, pp. 787 – 801, 2010.
- [11] H. Boessenkool, D. Abbink, C. J. M. Heemskerk, and F. C. T. Van der Helm, "Haptic shared control improves tele-operated task performance towards performance in direct control," in *World Haptics Conference (WHC)*, 2011 IEEE, 2011, pp. 433–438.

- [12] S. Nudehi, R. Mukherjee, and M. Ghodoussi, "A shared-control approach to haptic interface design for minimally invasive telesurgical training," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 4, pp. 588–592, Jul. 2005.
- [13] D. Powell and M. O'Malley, "Efficacy of shared-control guidance paradigms for robot-mediated training," in *World Haptics Conference (WHC), 2011 IEEE*, Jun. 2011, pp. 427–432.
- [14] D. Kragic, P. Marayong, M. Li, A. M. Okamura, and G. D. Hager, "Human-machine collaborative systems for microsurgical applications," *International Journal of Robotics Research*, vol. 24, pp. 731–741, 2005.
- [15] Y. Li, K. P. Tee, W. L. Chan, R. Yan, Y. Chua, and D. K. Limbu, "Role adaptation of human and robot in collaborative tasks," in *IEEE International Conference on Robotics and Automation*, 2015.
- [16] A. Mörtl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche, "The Role of Roles: Physical Cooperation between Humans and Robots," *The International Journal of Robotics Research*, Aug. 2012.
- [17] P. Évrard and A. Kheddar, "Homotopy switching model for dyad haptic interaction in physical collaborative tasks," in *EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*, Mar. 2009, pp. 45–50.
- [18] D. Aarno, S. Ekvall, and D. Kragic, "Adaptive virtual fixtures for machine-assisted teleoperation tasks," in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation, 2005. ICRA 2005*, 2005, pp. 897–903.
- [19] R. Groten, "Haptic human-robot collaboration: How to learn from human dyads," Ph.D. dissertation, Technische Universität München, 2011.
- [20] A. Kucukyilmaz, T. Sezgin, and C. Basdogan, "Intention Recognition for Dynamic Role Exchange in Haptic Collaboration," *IEEE Transactions on Haptics*, vol. 6, no. 1, pp. 58–68, 2013.
- [21] A. D. Dragan and S. S. Srinivasa, "A Policy Blending Formalism for Shared Control," *The International Journal of Robotics Research*, vol. 32, no. 7, 2013.
- [22] K. B. Reed and M. A. Peshkin, "Physical collaboration of human-human and human-robot teams," *Haptics, IEEE Transactions on*, vol. 1, no. 2, pp. 108–120, Jul. 2008.
- [23] P. Marayong, A. Bettini, and A. Okamura, "Effect of virtual fixture compliance on human-machine cooperative manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2002*, vol. 2, 2002, pp. 1089–1095 vol.2.
- [24] H. Boessenkool, D. Abbink, C. Heemskerk, F. van der Helm, and J. Wildenbeest, "A task-specific analysis of the benefit of haptic shared control during telemanipulation," *IEEE Transactions on Haptics*, vol. 6, no. 1, pp. 2–12, 2013.
- [25] J. R. Busemeyer and J. G. Johnson, "Computational models of decision making," *Blackwell handbook of judgment and decision making*, pp. 133–154, 2004.
- [26] C. Woodruff, L. Vu, K. Morgansen, and D. Tomlin, "Deterministic modeling and evaluation of decision-making dynamics in sequential two-alternative forced choice tasks," *Proceedings of the IEEE*, vol. 100, no. 3, pp. 734–750, Mar. 2012.
- [27] C. Gonzalez, "Cognitive science: An introduction," in *Catheter-Based Cardiovascular Interventions: A Knowledge-Based Approach*, P. Lanzer, Ed. Springer, 2013, ch. 6, pp. 61–67.
- [28] C. Gonzalez and J. Quesada, "Learning in dynamic decision making: The recognition process," *Computational & Mathematical Organization Theory*, vol. 9, no. 4, pp. 287–304, 2003.
- [29] D. M. Egelman, C. Person, and P. R. Montague, "A computational role for dopamine delivery in human decision-making," *Cognitive Neuroscience*, p. 623630, 1998.
- [30] A. Stewart, M. Cao, A. Nedic, D. Tomlin, and N. Leonard, "Towards Human&Robot teams: Model-based analysis of human decision making in two-alternative choice tasks with social feedback," *Proceedings of the IEEE*, vol. 100, no. 3, pp. 751–775, 2012.
- [31] J. G. Johnson and J. R. Busemeyer, "A unified computational modeling approach to decision making," in *Proceedings of the Seventh International Conference on Cognitive Modeling*, 2006.
- [32] S. Meisel, *Anticipatory Optimization for Dynamic Decision Making*, 1st ed., ser. Operations Research/Computer Science Interfaces Series. Springer, 2011, vol. 51.
- [33] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks," *Psychological Review*, vol. 113, no. 4, p. 700765, Oct. 2006.
- [34] M. Cao, A. Stewart, and N. Leonard, "Integrating human and robot decision-making dynamics with feedback: Models and convergence analysis," in *47th IEEE Conference on Decision and Control, 2008. CDC 2008*, Dec. 2008, pp. 1127–1132.
- [35] J. Corredor, J. Sofrony, and A. Peer, "Deciding on optimal assistance policies in haptic shared control tasks," in *IEEE International Conference on Robotics and Automation*, 2014.
- [36] H. Yu, M. Spenko, and S. Dubowsky, "An adaptive shared control system for an intelligent mobility aid for the elderly," *Autonomous Robots*, vol. 15, no. 1, pp. 53–66, 2003.
- [37] E. J. Veras, K. De Laurentis, and R. Dubey, "Design and implementation of visual-haptic assistive control system for virtual rehabilitation exercise and teleoperation manipulation," in *30th Annual International IEEE EMBS Conference*, 2008, pp. 4290–4293.
- [38] Y. Li, J. Huegel, V. Patoglu, and M. O'Malley, "Progressive shared control for training in virtual environments," in *World Haptics 2009*, 2009, pp. 332–337.
- [39] R. Bogacz, S. M. McClure, J. Li, J. D. Cohen, and P. R. Montague, "Short-term memory traces for action bias in human reinforcement learning," *Brain Research*, vol. 1153, p. 111121, 2007.
- [40] A. Bettini, P. Marayong, S. Lang, A. Okamura, and G. Hager, "Vision-assisted control for manipulation using virtual fixtures," *IEEE Transactions on Robotics*, vol. 20, no. 6, pp. 953–966, Dec. 2004.
- [41] A. Kron, G. Schmidt, H. Baier, M. Buss, F. Freyberger, J. Hoogen, P. Kammermeier, and G. Schmidt, "Distributed PC-based haptic, visual and acoustic telepresence system-experiments in virtual and remote environments," in *Virtual Reality, 1999. Proceedings*, IEEE, vol. 4, 1999, pp. 118–125.
- [42] D. Feth, B. A. Tran, R. Groten, A. Peer, and M. Buss, "Shared-control paradigms in multi-operator-single-robot teleoperation," in *Cognitive Systems Monographs*. Springer, 2009.
- [43] M. Jang, M.-S. Han, J.-h. Kim, and H.-S. Yang, "Dynamic time warping-based k-means clustering for accelerometer-based handwriting recognition," in *Developing Concepts in Applied Intelligence*. Springer Berlin Heidelberg, 2011, no. 363, pp. 21–26.



**Javier Corredor** received the bachelors degree in Design and Electronic Automation Engineering from the Universidad de la Salle and the Masters Degree (Meritorious mention) in Engineering - Industrial Automation from the Universidad Nacional de Colombia, where he is currently working towards Ph.D degree at the Department of Industrial and Systems Engineering. His research interest include robotics, the design of haptic assistances for shared control systems and human-robot interaction.



**Jorge Sofrony** awarded his Ph.D. from the university of Leicester, UK, in the area of control systems. His research was part of the GARTEUR AG-15 initiative and consisted in designing compensator augmentations for systems subject to sector bounded nonlinearities. The results and design techniques have been fight-tested in collaboration with DLR (Germany), the University of Leicester. He is currently Associate Lecturer at the Department of Mechanical Engineering and Mechatronics, Universidad Nacional de Colombia. His current interest are parallel robotics and control, path-planning and aerospace systems, with special focus on FDI/FTC.



**Angelika Peer** is currently Full Professor at the Bristol Robotics Laboratory, University of the West of England, Bristol, UK. Before she was senior researcher and lecturer at the Institute of Automatic Control Engineering and TUM-IAS Junior Fellow of the Institute of Advanced Studies of the Technical University of Munich, Germany. She received the Diploma Engineering degree in Electrical Engineering and Information Technology in 2004 and the Doctor of Engineering degree in 2008 from the same university. Her research interests include robotics, haptics, teleoperation, human-human and human-robot interaction as well as human motor control.