2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Deciding on Optimal Assistance Policies in Haptic Shared Control Tasks

Javier Corredor[1], Jorge Sofrony[1] and Angelika Peer[2]

***Abstract*— This paper presents a haptic assistant that enhances task performance and human-machine interaction via a gain-scheduled impedance controller. The assistance strategy proposed builds on decision-making studies and models first proposed in the field of cognitive science and combines these models with a gain-scheduled impedance control technique in order to enhance human machine interaction in a tracking task with environmental uncertainties. This paper explores the Drift-Diffusion Model as decision making model and proposes an adaptive impedance control strategy that enhances both, task performance and human-machine interaction.**

I. INTRODUCTION

Bilateral teleoperation systems have gained importance in application fields like medical robotics, micro/macro manipulation, manipulation of robots in hazardous environments as well as remote control of unmanned vehicles, space and mobile robots [1]. Haptic teleoperation systems come with challenges in terms of human-machine interaction (HMI) and closed-loop control. Although the main underlying control challenges (i.e. nonlinearities and uncertainties in the robot manipulator model, time-delays in the communication channel, varying environment and human impedances) have been widely addressed in literature (see [2] for an overview), the interest in HMI has grown in recent years [3]–[5]. Nonetheless, most of the developments in this field have focussed on modeling the human intention, with human-machine cooperation receiving a considerable amount of attention recently [6], [7].

Cooperative control schemes require humans and robots to collaborate as *peers*, implying a change from a strategy where the human is a supervisor, to another, where the robot is allowed to make its own decisions and to *collaborate* with the operator. Hence, it is necessary to consider an interaction strategy where the operator's and robot's authority are adjustable over the task.

Haptic assistants implementing shared control and changing authority have been applied with techniques ranging from a constant level of assistance [8], to assistances with switching [9], linear [10] and nonlinear [11] adapting mechanisms. Nonetheless, these controllers tend to have scheduling dynamics that may be not very intuitive for the operator and hence, they may present poor performance or unwanted behavior in unknown environments. In order to produce a scheduling/switching strategy that results naturally intuitive to the operator, [12] first proposed to incorporate a *Decision-Making Model* to aid the human operator make suitable choices in a cooperative robotic-human foraging task; in this paper we propose to use the *Drift-Diffusion* (DD) model as the main scheduling operator in a haptic human-robot collaboration task, and presume that this will result in an intuitive assistance strategy that may enhance collaboration.

DD models describe the decision making mechanism in humans when confronted with a Two-Alternative Force Choice (TAFC) task; the model is constructed based on findings of behavioral studies performed in the field of cognitive science [13]. The model describes the decision making mechanism presented by humans as a process in which based on past decisions predictions about the success of future decisions are made, and the decision criteria is continuously adjusted in order to maximize the reward obtained throughout the task.

The paper is organized as follows: First we present the general control structure and discuss the gain scheduled admittance control strategy. Then, we present the decision making model and propose a suitable scheduling strategy in the context of a TAFC decision making problem. This is followed by some experimental results. The paper closes by giving some concluding remarks.

II. PROBLEM STATEMENT

In teleoperation it is desirable for the human to maintain control over the task, for example, when an obstacle or any other environment uncertainty, unknown to the assistance (but known to the operator) is present. On the other hand, when the assistance has full knowledge of the environment, we wish to give more authority to the remote robot in order to enhance task performance. In fact, if the task is static and is developed in a structured environment, the robot can handle the execution autonomously. We consider the problem where the environment can be unstructured or partially structured (i.e. composed of clear tracking objectives and unmodeled obstacles), making it difficult or even impossible to precisely model the entire task *a priori*.

Both objectives (e.g. trajectory tracking and obstacle avoidance) are in contradiction since good task performance requires tight control, while unrestricted movements require a less aggressive control strategy. This trade-off was explored in our previous work [14], and we found that in case of a continuous, multi-criteria assistance policy, unclear (or unnatural) haptic feedback signals may be sent to the human, hence the meaning of these signals may not be intuitive for

[1]Javier Corredor and Jorge Sofrony are with the Faculty Engineering, Universidad Nacional de Colombia, {jcorredorc, jsofronye}@unal.edu.co

[2]Angelika Peer is with the Institute of Automatic Control Engineering, Technische Universität München, angelika.peer@tum.de

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 are supported by the National PhD Support Program, Instituto Colombiano para el desarrollo de la Ciencia y la Tecnología (COLCIENCIAS), Colombia.

978-1-4799-3684-7/14/$31.00 ©2014 IEEE

the operator. We also found that a switching policy between high and low haptic assistance based on a predefined and time-invariant decision criterion, allows to achieve a good trade-off between task performance and agreement of human and haptic assistant.

In the literature, most assistance mechanisms move between levels of assistance that are thought to improve some metrics involved in the task [15], [14], hence an inherent difficulty is to establish suitable metrics and scheduling strategies such that the haptic feedback results in a natural assistance for the operator [14]. In this paper we propose to consider the allocation of control authority as a decision making process in a TAFC problem setup, where our choices are whether to assist the human to improve task performance or interaction performance. In the proposed assistance scheme, we use ideas borrowed from behavioral studies performed in the field of cognitive science to adapt the level of assistance (authority) by varying the stiffness of the haptic device.

III. Control System Architecture

The teleoperation setup presented in this paper is standard and considers two robot manipulators: the master (haptic) device and the slave device. We assume that an exact linearization control scheme is present at the master and slave sides, thus the plant considered in Fig. 1 is assumed to be linear. This (linear) plant is later augmented with a PD (internal loop) controller in order to give nominal position tracking dynamics. If we want to imprint some type of "reference" dynamics to the system (for example if we want to vary the system's apparent stiffness), an additional block must be introduced. The additional filter Y implements the desired dynamics and outputs a reference position (see Fig. 1). This position-based admittance controller will be used later as the main mechanism to vary the level of assistance perceived by the human operator.

Stability of the overall teleoperation architecture may be concluded from the following two steps: the first step is concerned with showing that the master side is stable and strictly passive (SP) for any (bounded) variation of the scheduling parameter suggested, hence we have no restrictions on the velocity of the scheduling parameter (only the mild assumption of boundedness); the second step uses the previous fact to conclude that under mild and general assumptions, the teleoperation system is stable since we have the feedback interconnection of two SP components (see [16]). In general we assume that the robot manipulators at the master and slave sides have an exact model, that interaction with the environment is strictly passive (SP) and that the communication channel has no time delays, hence we are only concerned with the former step.

A. Variable gain admittance control

Note that the master (or slave) device may be modeled as a mass-damper system [17], [18]. If we assume that the low-level position controller has fast and well damped dynamics, then a position-based admittance controller may be used to imprint some reference dynamics to the master device and

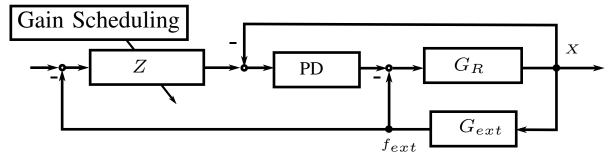

Fig. 1. Position-based admittance control: varying admittance (Y), G_R haptic device and G_{ext} operator dynamics.

give the human the sensation that the system is more or less stiff or damped. By varying either the stiffness (k) or damping (b) it is possible to change the system dynamics, which can translate into a more or less assistive control strategy; hereafter we will assume that only k is varied.

Since the internal control loop is considered to be SP, the interconnection of the reference admittance and this internal loop is SP if Y is SP. Nonetheless, since the reference admittance is a parameter varying system, we must show this system is in fact SP. We will now restrict ourselves to the analysis of the dynamics of Y.

The state space model for the admittance controller has as states the position (x_1) and the velocity (x_2). The relating state space equations are:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k(t)}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u(t), \quad (1)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \quad (2)$$

First we must ensure that the time-varying filter, as made explicit by the term $k(t)$ in the A-matrix in (1), is strictly passive for any bounded variation $k(t)$. Assume that the parameter variation is bounded, i.e. $k(t) \in [k_1, k_2]$, and define A_1, A_2 as the A-matrix of the state-space system when $k(t) = k_1$ and $k(t) = k_2$, respectively. The system may now be described as an affine linear parameter varying system (a-LPV), which can be described by its convex polytopic set. This set is spanned by its vertices computed at the upper and lower limit of the uncertain parameter k,

$$\dot{\boldsymbol{x}} = \sum_{i=1}^{m} \rho_i A_i, \qquad \sum_{i=1}^{m} \rho_i = 1, \quad (3)$$

where $m = 2$. For a-LPV systems, stability of the autonomous system (3) may be guaranteed if there exists a Common Quadratic Lyapunov Function (CQLF) to all the subsystems such that

$$A_i^T P + P A_i < 0 \quad \forall i. \quad (4)$$

If this is the case, stability of (3) is guaranteed under *any* variation of the parameter [19] (i.e. the velocity of $k(t)$ is not restricted). If $m = 2$ in (3), the existence of a CQLF is guaranteed under certain conditions as detailed in the following theorem.

Theorem 1: [20], [21]: Given two Hurwitz matrices A_1, $A_2 \in \mathbb{R}^{n \times n}$ such that $\mathrm{rank}(A2 - A1) = 1$, a necessary and sufficient condition for the existence of a CQLF is that the matrix product $(A_1 A_2)$ has no negative real eigenvalues.

Notice that in the haptic assistance scenario, the difference $(A_1 - A_2)$ is always rank $= 1$ by construction. The eigenvalues for $(A_1 A_2)$ are,

$$\lambda_{1,2} = -\frac{k_1 m + k_2 m - b^2 \pm \sqrt{\eta}}{2m^2} \quad \text{with} \tag{5}$$

$\eta = b^4 - 2\,b^2\,\mathrm{k}_1\,m - 2\,b^2\,\mathrm{k}_2\,m + {\mathrm{k}_1}^2\,m^2 - 2\,\mathrm{k}_1\,\mathrm{k}_2\,m^2 + {\mathrm{k}_2}^2\,m^2.$

The existence of a CQLF for (4) is guaranteed if:

$$\eta < 0 \qquad k_1 m + k_2 m - b^2 < 0. \tag{6}$$

Finally, we may conclude that the interconnection presented in Fig. 1 is stable if conditions (6) are satisfied.

B. Teleoperation control scheme

Consider a tracking task where the desired path is known to the haptic assistant.

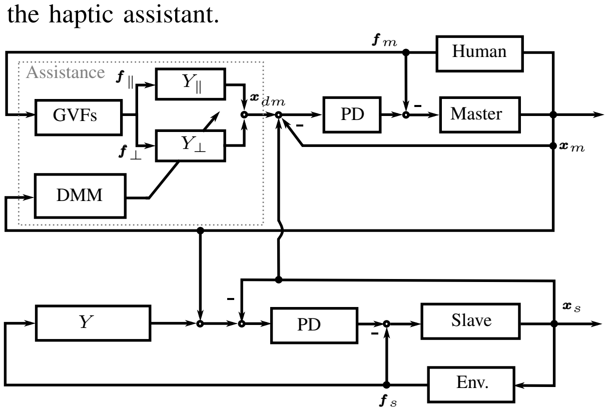

Fig. 2. Scheme of the decision making based assistance, the assistance is composed by the decision making model (DMM) and the guidance virtual fixture (GVFs). The control scheme is a variable admittance (Y) and a PD controller

The force measured at the tip of the master can be decomposed as $\boldsymbol{f}_m = \boldsymbol{f}_{\|} + \boldsymbol{f}_{\perp}$, where $\boldsymbol{f}_{\|}, \boldsymbol{f}_{\perp}$ are the forces in the parallel and perpendicular direction of the desired path, respectively. Without any loss of generality, we will assume that a time-varying admittance block will only be present in the perpendicular direction, since we want to help the operator to stay on the path and not to move along it. The admittance in the parallel direction ($Y_{\|}$) is given by

$$\boldsymbol{f}_{\|} = m\ddot{\boldsymbol{x}}_{\|} + b_{\|}\dot{\boldsymbol{x}}_{\|}, \tag{7}$$

where m is the mass and $b_{\|}$ the damping coefficient for the parallel admittance. The admittance in the perpendicular direction ($Y_{\perp}$) is given by

$$\boldsymbol{f}_{\perp} = m\ddot{\boldsymbol{x}}_{\perp} + b_{\perp}\dot{\boldsymbol{x}}_{\perp} + k_{\perp}\boldsymbol{x}_{\perp}, \tag{8}$$

where $b_{\perp}$ and $k_{\perp}$ are the damping coefficient and the stiffness, respectively. The desired position of the master is then given by

$$\boldsymbol{x}_{dm} = \boldsymbol{x}_{\|} + \boldsymbol{x}_{\perp}. \tag{9}$$

Note that the assistance must make a decision about the stiffness ($k_{\perp}$) that is suitable to improve the task or the interaction performance, hence the decision making model presented in the next section will adapt the level of assistance (see $Y_{\perp}$ in Fig. 2).

IV. Decision making models

In cognitive science it was found that humans make decisions based on previous rewards obtained on past decisions, and that they estimate future rewards in order to maximize their total intake (see e.g. [22]). In this paper, we take inspiration from this finding to implement a similar behavior to our haptic assistance, where the reward may be calculated based on some measure of the performance, including the dynamics of the interaction between the human and the assistance. In this paper we present a haptic assistance scheme where the reward is calculated based on the tracking error as task performance measure (T_P), and the interaction forces as interaction performance measure (I_P).

In cognitive science, the *two-alternative forced-choice (TAFC)* task was proposed as an experiment to explore human decision making mechanisms [13]. In this task the human is confronted with a two alternative, sequential choice problem. The alternatives in the TAFC task are intended to induce dynamics where the human has to exploit available or new resources. Assume that only two choices are available (i.e. namely T and I), and that the reward $r(t)$ is given by

$$r(t) = \begin{cases} r_I(I_P) & \text{if } z(t) = I, \\ r_T(T_P) & \text{if } z(t) = T, \end{cases} \tag{10}$$

where $z(t) \in [T, I]$ is the decision made at time t. The functions r_T and r_I represent the reward calculated for each choice, where T_P and I_P are the established performance measures, and:

$$r_I = k_I I_P + I_{P0} \qquad r_T \;= k_T T_P + T_{P0}, \tag{11}$$

where k_T and k_I are the ratio of the change (slope), I_{P0} and T_{P0} are the reward at the lowest and highest performance, respectively. Due to their contradicting nature, both choices may be represented as linear functions with contradicting slopes as depicted in Fig. 3 (see [14]).

Now it is necessary to find a method that allows the appropriate selection of $z(t)$ such that the total intake is maximized. For this purpose, the Drift Diffusion Model (DD) proposed in cognitive science literature will be used.

A. Drift Diffusion Model

The DD model is a reinforcement learning model, in which evidence in favor of one choice is integrated until a predetermined threshold is reached; the *soft-max* model is used as the DD model strategy such that the probability of

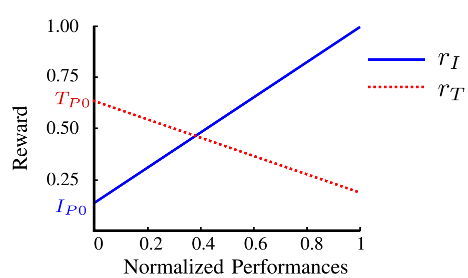

Fig. 3. The matching shoulder task with linear reward schedules

choosing I over T (or vice versa) is given by

$$p_I(t+1) = \frac{1}{1+\exp^{-\mu(w_I - w_T)}}, \tag{12}$$

where μ is the slope of the sigmoidal function, and w_I and w_T are the *evidence* in favor of (I) and (T) respectively.

Remark 1: Larger μ represents a high slope in the sigmoidal function (12), this means that there is more certainty in the decision making process, hence the commutation between gains is faster.

In order to take performance information into account, an update rule for the evidence in favor of each choice is proposed as

$$\begin{aligned} w_z(t+1) &= w_z(t) + \lambda(r_z(t) - w_z(t)), \\ w_{\bar{z}}(t+1) &= \lambda r_{\bar{z}}(t), \end{aligned} \tag{13}$$

where $z \in [I, T]$ represents the decision just made, $r_z(t)$ is the reward of decision z, $\lambda \in [0, 1]$ is the learning rate and the symbol $\bar{\cdot}$ denotes the "not" operator. This is inspired by the reinforcement rule presented in [22], where the evidence for the decision just made (w_z) is accumulated after each decision and the weight is updated based on the difference between the actual reward r_z and the expected reward w_z. The evidence of the option that is not chosen ($\bar{z}$) is updated without any learning rule, just with the previous reward obtained. The parameter λ, in the $\bar{z}$ choice, scales the reward to maintain no preference over any option when interaction commences.

Remark 2: The accumulated evidence can be seen as the memory of the model of previous rewards. Putting more weight on memory means that the model takes into account history more when making a decision. The parameter λ represents the memory of the DDM. If λ is low, the model takes into account history more, this is, memory (and vice-versa). The decision making rule proposed is given by

$$z(t+1) = \begin{cases} I & \text{if } p_I > 0.5 \\ T & \text{otherwise.} \end{cases} \tag{14}$$

V. DDM-BASED HAPTIC ASSISTANCE

The setup proposed consists of tracking a predefined path combined with unmodeled obstacles that the operator must evade (see Fig. 4).

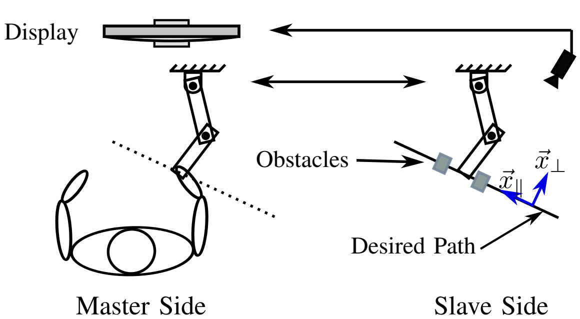

Fig. 4. Teleoperation task

Performance measures are calculated in the perpendicular direction (see Section III-B) of the path over an observation window of N samples. The tracking performance is measured via the position error between the desired path and the current position of the master; the normalized version of the tracking performance measure is given by

$$T_{P,N} = \frac{mean_N(\|e\|)}{max(\|e\|)},$$

where $T_{P,N}$ is the normalization of the task performance and e is the position error, where $max(\|e\|)$ is calculated for each decision window.

Internal forces occur if two partners push or pull in different directions [14] and may be used as an interaction measure. These forces are wasted effort from a physical point of view because they do not contribute to the movement of the object, but still provide important information on the haptic interaction and negotiation strategy [7]. The (haptic) interaction performance is defined as the agreement measurement between the operator and the master device. A measure of these internal forces is defined as in [7]

$$f_i = \begin{cases} f_\perp & \text{if } sign(f_\perp) \neq sign(f_{a\perp}) \wedge \|f_\perp\| \leq \|f_{a\perp}\| \\ -f_{a_\perp} & \text{if } sign(f_\perp) \neq sign(f_{a_\perp}) \wedge \|f_\perp\| > \|f_{a_\perp}\| \\ 0 & \text{if otherwise} \end{cases}, \tag{15}$$

where f_i represents the internal forces, $f_\perp$ is the force exerted by the human and f_a is the force exerted by the assistance. All forces are measured in the perpendicular direction of the desired path. The interaction performance is finally defined by

$$I_{P,N} = 1 - f_{i,N}, \qquad f_{i,N} = \frac{mean_N(\|f_i\|)}{max(\|f_i\|)}, \tag{16}$$

where $I_{P,N}$ is the interaction performance and is referred to as agreement measure; $f_{i,N}$ is the normalized version of the internal forces. The level of the haptic assistance depends on the probability that a certain choice improves the total system reward. Define the scheduling parameter function $\alpha \in [0\ 1]$ as

$$\alpha(p_I) = \frac{1}{2} + \frac{1}{2}\tanh\left(\frac{p_I - \phi}{\varphi}\right) \tag{17}$$

where ϕ and φ are user defined parameters; ϕ is the switch point and φ is the smoothing level. The function α maps the probability (p_I) to the level of haptic assistance via a linear homotopy given by

$$k = \alpha\ k_{low} + (1-\alpha)\ k_{high} \tag{18}$$

where k is the assistance level, k_{high} and k_{low} are the maximum and minimum values of the assistance, respectively.

VI. EXPERIMENTAL RESULTS

The data presented was recorded in a real teleoperation system consisting of two admittance-type haptic feedback devices with 4-DOF (see [23] for details). The human operator was asked to move along a predefined path known to the haptic assistant, but had to avoid unmodeled obstacles (see Fig. 4). The parameters for the target dynamics are:

$$m_\perp = 5, \quad b_\perp \quad = 200, \quad k1_\perp = 1000, \quad k2_\perp \quad = 100.$$

Conditions (6) are satisfied, and the eigenvalues of $(A1A2)$ are $\sigma_{1,2} = [2.9 \quad 1377.1]$, therefore stability is guaranteed for (bounded) parameter variations. To observe the effect of changing some user defined parameters, different values of the parameters μ and matching shoulder structure are considered. The λ parameter is selected to be constant in all the experiments to reduce the number of parameters to be analyzed simultaneously.

When the operator moves off the path, the position error ($T_{P,N}$) increases and the agreement ($I_{P,N}$) decreases (cf. Perf. in Fig. 5 and 6). Marked time windows indicate phases in which the operator moved off the path.

In general, the μ-parameter adjusts the degree of certainty to improve a certain performance index. In the experiment conducted, having large certainty (high probability) to improve interaction performance, means that small movements outside the path are associated with a fast change in the levels of assistance, (see Fig. 5 at 5-7 and 9-11 sec.). This behavior may resemble the proposed strategy in [24], which assist the operator when it lies in a given region and no assistance is provided outside this region. Even though this type of fast changing dynamics can be desirable when the operator is required to maintain control of the robot in light of sudden changes in the remote environment, quick changes in the level of assistance are felt as unnatural by the operator, hence decreasing the perceived comfort. On the other hand, if μ is gradually decreased to lower values (see Fig. 6), the operator does not experiment sudden changes in the level of assistance, and the interaction may be perceived as more natural.

Different scheduling dynamics may also be achieved by changing the reward structure to assist one objective with more preference. For example in telesurgery, it might be necessary to restrict the movement of the operator to a given path. In other tasks path tracking may be less important, for example, when it is necessary that the operator avoids

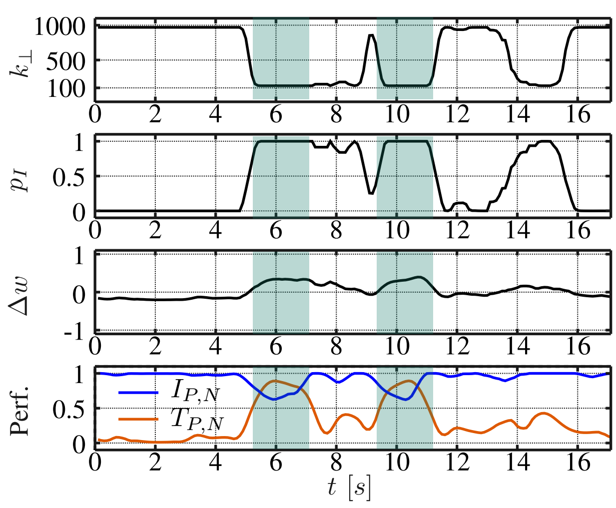

Fig. 5. Drift-Diffusion model-based assistance response in the perpendicular direction of the desired path, for $\lambda = 0.7$ and $\mu = 100$

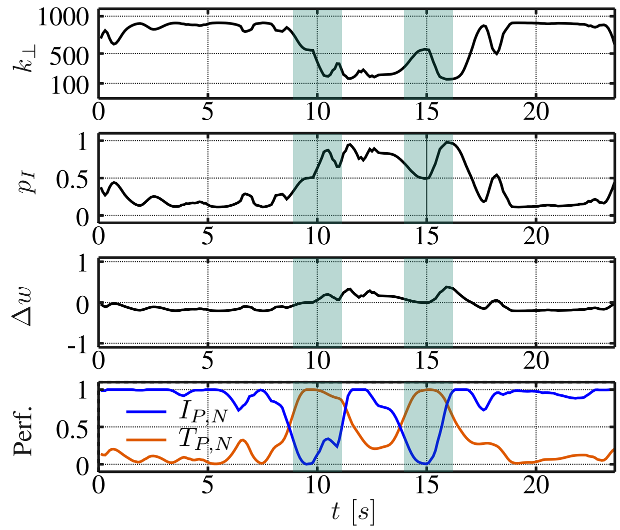

Fig. 6. Drift-Diffusion model response in the perpendicular direction of the desired path, for $\lambda = 0.7$ and $\mu = 10$

obstacles. These situations may be modeled by the reward structures (r_I, r_T). The parameters k_T, k_I, T_{P0} and I_{P0} (see (11)) adjust the preference to assist the operator for the task or interaction.

The preference to assist the operator predominantly in tracking may be achieved by considering $k_T = -1, k_I = 0.5, T_{P0} = 1, I_{P0} = 0$. The maximum reward for interaction is half of the one set in the previous experiment, giving more importance to task performance rather than to interaction performance. Likewise to weight the interaction performance, the reward for the task performance can be reduced; for example, considering $k_T = -0.5, k_I = 1, T_{P0} = 0.5$ and $I_{P0} = 0$.

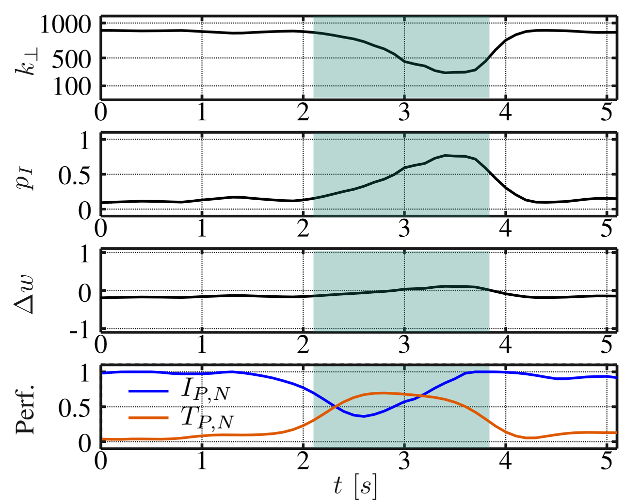

Fig. 7. Response for the DDM-based assistance with task preference (T_P), $k_T = -1, k_I = 0.5, c_T = 1$ and $c_I = 0$

When task performance is preferred over interaction performance, the evidence Δw to choose a stiffness that improves interaction performance is low, therefore the probabil-

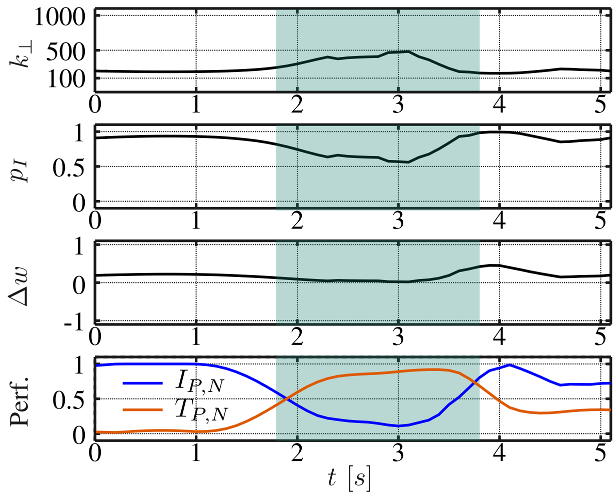

Fig. 8. Response for the DDM-based assistance with interaction preference (I_P), $k_T = -0.5, k_I = 1, c_T = 0.5$ and $c_I = 0$

ity p_I is generally low (Fig. 7). In contrast, when interaction is preferred over task performance, the evidence Δw to assist in the interaction is generally high, so the probability to assist the operator to improve the interaction is also generally high (Fig. 8).

When the task is preferred 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 (mark in Fig. 7), the evidence to assist in improving interaction performance increases and the probability p_I also increases slowly. The user feels a high level of assistance when he/she moves on the path and the stiffness slowly decreases when moving off the path, but generally the stiffness level remains quite high in favor of improving task performance.

When interaction is preferred and the user moves on the path, the agreement and the tracking error accumulate evidence to assist in improving interaction performance. On the other hand, when the user moves off the path (mark in Fig. 8), the agreement decreases and the probability p_I reduces because the evidence Δw is accumulated at slow rate. The user feels a low stiffness when he/she moves on the path while the stiffness slightly increases when the user moves off the path, but generally the stiffness level remains quite low in favor of improving interaction performance.

VII. Conclusions

Our results showed that the DD model is a suitable scheduling strategy for haptic shared control, where the decision making process can be influenced via the parameters of the reward functions. The DD model consists of two parts: one that models the dynamics of the decision-making process (soft-max rule) and another that models different dynamics of the interaction between the assistance and the operator.

Cao et.al. [12] have proposed adapting the slope of the reward structures online in the DD model to assist the operator in making optimal decisions. Future work will target 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.

References

[1] G. Niemeyer, C. Preusche, and G. Hirzinger, "Telerobotics," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, 2008, pp. 741–757.

[2] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, Dec. 2006.

[3] R. Aracil, M. Buss, S. Cobos, M. Ferre, S. Hirche, M. Kuschel, and A. Peer, "The human role in telerobotics," *Advances in Telerobotics*, 2007.

[4] D. Powell and M. O'Malley, "The task-dependent efficacy of shared-control haptic guidance paradigms," *IEEE Transactions on Haptics*, vol. 5, no. 3, pp. 208–219, 2012.

[5] A. Kheddar, "Human-robot haptic joint actions is an equal control-sharing approach possible?" 2011, pp. 268–273.

[6] J.-M. Hoc, "From human-machine interaction to human-machine cooperation," *Ergonomics*, vol. 43, no. 7, pp. 833–843, 2000, PMID: 10929820.

[7] R. Groten, "Haptic human-robot collaboration: How to learn from human dyads," Ph.D. dissertation, Technische Universität München, 2011.

[8] 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," 2011, pp. 433–438.

[9] Y. Li, V. Patoglu, and M. K. . O'Malley, "Negative efficacy of fixed gain error reducing shared control for training in virtual environments," *ACM Transactions on Applied Perception*, vol. 6, pp. 1–21, 2009.

[10] P. Evrard and A. Kheddar, "Homotopy switching model for dyad haptic interaction in physical collaborative tasks," 2009, pp. 45 –50.

[11] H. Yu, M. Spenko, and S. Dubowsky, "An adaptive shared control system for an intelligent mobility aid for the elderly," in *Autonomous Robots*, vol. 15, 2003, pp. 53–66.

[12] M. Cao, A. Stewart, and N. Leonard, "Integrating human and robot decision-making dynamics with feedback: Models and convergence analysis," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, Dec. 2008, pp. 1127 –1132.

[13] P. Montague and G. S. Berns, "Neural economics and the biological substrates of valuation," *Neuron*, vol. 36, no. 2, pp. 265–284, 2002.

[14] 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.

[15] S. Oguz, A. Kucukyilmaz, T. Sezgin, and C. Basdogan, "Haptic negotiation and role exchange for collaboration in virtual environments," in *Haptics Symposium, 2010 IEEE*, Mar. 2010, pp. 371 –378.

[16] H. Khalil, *Nonlinear Systems*. New Jersey: Prentice Hall, 1996.

[17] Y. Yokokohji and T. Yoshikawa, "Bilateral control of master-slave manipulators for ideal kinesthetic coupling-formulation and experiment," in , *1992 IEEE International Conference on Robotics and Automation, 1992. Proceedings*. IEEE, May 1992, pp. 849–858 vol.1.

[18] A. Peer and M. Buss, "Robust stability analysis of a bilateral teleoperation system using the parameter space approach." IEEE, Sept. 2008, pp. 2350–2356.

[19] S. Boyd, L. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994, vol. 1.

[20] H. Lin and P. J. Antsaklis, "Stability and stabilizability of switched linear systems: A survey of recent results," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 308–322, Feb. 2009.

[21] R. N. Shorten, O. Mason, F. O'Cairbre, and P. Curran, "A unifying framework for the siso circle criterion and other quadratic stability criteria," *Int. Journal of Control*, vol. 77, no. 1, pp. 1–8, 2004.

[22] 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, pp. 111–121, 2007.

[23] M. Buss and G. Schmidt, "Control problems in multimodal telepresence systems." London: Springer-Verlag, 1999, pp. 65–101.

[24] 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.