Adaptive assistance-based on decision-making models for telerobotics systems

A thesis
submitted to the Faculty of Engineering
of the Universidad Nacional de Colombia
by

Javier Adolfo Corredor Camargo

In partial fulfilment of the requirements
for the degree of
Doctorate in Engineering - Computing and Systems

PhD. Jorge Sofrony Esmeral

Ap. 2016

Resumen

Esta tesis propone una nueva estrategia de asistencia háptica en la interacción humano-robot. Dado que el humano es el elemento fundamental del sistema, es necesario proponer estrategias que se adapten a su comportamiento, además de garantizar un mejoramiento del desempeño en la tarea. El inconveniente surge cuando se requiere asistir al operador en mejorar el desempeño de la tarea y permitir al usuario total control de la tarea cuando sea necesario, desviándose del plan original con el objetivo de abordar situaciones imprevistas.

Desde una perspectiva enfocada en el control, se debe resolver el compromiso existente entre proveer un alto nivel de asistencia para mejorar el desempeño de la tarea y un bajo nivel de asistencia para permitir al operador desviarse del plan pre-programado (original). Se propone entonces incorporar en la asistencia háptica un mecanismo de toma de decisiones usado por los humanos en tareas básicas de decisión entre dos alternativas. Este mecanismo de decisión se incorporar como el método de selección de parámetros en un controlador adaptativo de estructura fija (i.e. un controlador de impedencia/admitancia de parámetros variables).

Los resultados experimentales demuestran que el modelo de toma de decisión, i.e. el modelo *drift-diffusion* modificado, permite asignar el nivel de autonomía de una forma que resulta intuitiva para el usuario y mejora el desempeño en la tarea. Además la estrategia de asistencia basada en modelos de toma de decisión proporciona un mecanismo de sintonización para resolver diferentes requerimientos de la tarea, lo cual es importante en entornos no estructurados.

Dado el número de parámetros configurables presentes en la asistencia, la etapa experimental expone la función de cada uno de estos parámetros. Se realizó un experimento con usuarios en un entorno de teleoperación donde se evalúa estadísticamente el comportamiento de la asistencia en entornos parcialmente estructurados y se compara con la asistencia proporcionada por un experto humano, la cual puede ser considerada como la asistencia adaptativa nominal.

Palabras clave: *Telerobotica, teleoperación, control compartido, háptica, toma de decisión, drift-diffusion, WSLS*

Abstract

This thesis proposes a novel haptic assistance method for human-robot interaction. Since the human is the main element of the system, it is necessary to propose strategies that adapt the robot's dynamics to the human behavior, while guaranteeing an improvement in task performance. The main issue arises when the assistance must chose between assisting the operator to improve task performance or allowing the user to have full authority over the task when necessary, allowing him/her to deviate from the original plan in order to handle unforeseen situations.

From a control systems' perspective, the assistance has to solve the trade-off between high assistance levels to improve task performance and low assistance level to allow the user to deviate from the preprogrammed (original) plan. The main results of this work incorporate into the haptic assistance a human-like decision-making mechanism used in two-alternative force choice tasks. Our experimental results show that the *drift-diffusion*, which is a decision-making model proposed in the cognitive area, allocates control authority in a way that is intuitive for the user.

The the proposed assistance provides a tunable (decision-making) mechanism that is capable of fulfilling different task requirements, which is an important when dealing with unstructured environments. Given the number of configurable parameters in the assistance mechanism, the experimental procedure exposes the effects of changing them. A user study in a telerobotic scenario was performed to evaluate the behavior of the assistance in a partially structured environment; the proposed assistance is compared to the assistance provided by a human expert, which may be considered as the nominal adaptive assistance.

keywords: *Teleoperation, telerobotics, shared control, haptics, decision-making, drift-diffusion, WSLS, TAFC*

Acknowledgements

I would like to thank my advisor Dr. Jorge Sofrony. Without his kind support this thesis would have not been possible. I would also like to thank Dr. Angelika Peer, who gave me the chance to carry out my research at the Institute of Automatic Control Engineering (LSR) of the Technische Universität München. I am also thankful to her for reading my reports, commenting on my views and helping me understand and enrich this thesis.
My sincere thanks also goes to Ernesto Córdoba, who supported my application to the National PhD Support Program at the *Instituto Colombiano para el desarrollo de la Ciencia y la Tecnología (COLCIENCIAS)*, Colombia. Furthermore, I want to thank all my colleagues who supported this work at the Universidad Nacional de Colombia and the Technische Universität München.
Last but not least, I am deeply grateful to my wife Carolina and all my family for their never-ending support in all my endeavors.

Contents

List of Figures

List of Tables

1. Introduction

The creation of Human-Robot teams give raise to new challenges in the field of robotics. Robots may support humans when performimg tasks with high physical or mental load; they also give additional capabilities to human operators, allowing then to perform activities that otherwise s/he could have not handled by themselves without the robotic interfaces. This new scenario (i.e. when humans and machines interact closely) requires—besides ensuring human safety—of some strategy to provide the robot with mechanisms that allow it to interact "naturally" with the user; to achieve this goal the robot needs to *make decisions* in order to assist humans in uncertain environments. Robotics has shifted from a paradigm where the robot is physically separated from the human, to one in which appropriate human-robot interaction is at the foreground.

Telerobotic systems are one of the first applications in which human-robot interaction played a fundamental role. Telerobotic systems allow a human operator to perform tasks remotely. Ideally this should permit the operator to perform the task as if s/he conducted it directly without any intermediate systems involved. Haptic feedback—which refers to the use of tactile and/or force cues—provides information that increases the sense of presence at the remote site; this information can be augmented with cues that guide the user or inform him/her about different task stages. Telerobotic 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 [Niemeyer et al., 2008].

Although human operators have the ability to abstract new information, learn and make decisions in unexpected situations, when they perform a task using telerobotic 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 (i.e. restricted workspace or limited number of degrees of freedom (DoF)) that complicate the interaction.

A way of reducing the effect of the aforementioned limitations is to provide assistance to the user performing the task e.g. "guiding" user movements towards a successfull task execution. When unplanned situations occur, many challenges may arise, one of them being the construction of an adaptation mechanism which allows varying the level of the provided assistance between completely human-driven task execution and purely autonomous operation [Passenberg et al., 2013].

In teleoperation systems it is desirable that the operator maintains control over the task if any uncertainty occurs; uncertainties may be defined as any event unknown by the assistance, but known by the operator. The assistance mechanism can properly react to situations for which it was designed i.e. programmed actions which are known to improve certain performance criteria under structured environments. From a control systems' viewpoint, there exists a trade-off between a good tracking (task) performance, which requires tight control, and unrestricted human movements, which requires a less aggressive control strategy.

To date, previous studies have reported this trade-off [Abbink et al., 2012, Mörtl et al., 2012, Passenberg et al., 2013], and several approaches have been proposed to address it. For instance, techniques ranging from constant assistance levels [Boessenkool et al., 2011], to assistance with switching dynamics [Li et al., 2009]; linear [Evrard and Kheddar, 2009b] and nonlinear [Yu et al., 2003] adapting mechanisms. Although such techniques have proven to be somewhat successful, 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 unstructured (or partially structured) environments. Recently, game theory ideas have been included into tele robotic assis-

tance mechanisms [Li et al., 2015]; for instance, humans were capable of recognizing game-theory negotiation strategies through haptic feedback [Oguz et al., 2012]. Another example, although no haptic feedback is included, is the inclusion of a *dynamic decision-making model* to aid the human operator in order to make suitable choices in a cooperative robot-human foraging task [Cao et al., 2008].

Although modeling of the decision making process is an active research area, incorporating such processes as part of a negotiation strategy in teleoperation systems in order to provide more natural human-robot interaction is on its early attempts. Efforts are focussed on incorporating a dynamic decision-making model into the haptic assistance mechanism mainly because in complex tasks with unstructured environments, robots must make decisions which generally require solving nonlinear—or/and stochastic—optimization problems which in many cases are intractable [Cao et al., 2010].

Although it is possible to determine suboptimal decision making mechanisms, the designer cannot always guarantee an adequate response of the system due to the diversity of unexpected scenarios and the unknown parameters of the operator. As noted by Busemeyer et.al. [Busemeyer and Johnson, 2004], following only a utility based approach (as in classical optimization or classical decision theory) in decision-making may become complex. Thus, understanding the fundamentals of the decision-making mechanism in which the the utility representations are based may become simpler and produce a more coherent representation.

This is the case in real-applications of haptic assistance that must consider high uncertainty, complexity, and rapidly changing environments. In fact, the assistance takes advantage of additional human abilities such as reacting to unexpected situations, recognizing patterns and applying it to new situations (learning) in order to perform complex tasks in partially or unstructured environments. On the other hand, in these environments, the operator can be overloaded with information that may prevent him/her from making the correct decision; this because the operator omits key information or the information is not clearly available to the operator. In fact, the human decision-making process can be benefit from fast processing of large amounts of information that the robot can handle [Cao et al., 2010]. Dynamic decision making explores the human (and animals) need for speed and accuracy in trade-off decision making situations that exert strong evolutionary influences, thereby optimizing mechanisms [Bogacz et al., 2006].

Teleoperation tasks, in particular those in which movements are performed in free space, are benefited by focusing on shared control strategies, rather than by focusing on improving the transparency of the system [Boessenkool et al., 2013]. The influence of the haptic assistances has been experimentally validated in [Nitsch et al., 2012, Gunn et al., 2009, Powell and O'Malley, 2012].

Main contribution

This thesis proposes the allocation of control authority as a decision making process in which the choices consist of: (i) assisting the human to improve task performance via high control gains or (ii) giving the full control to the user via a low assistance level. In the proposed assistance scheme, ideas borrowed from behavioral and neuronal experiments performed in the field of cognitive science are used to adapt the assistance level (authority) by varying the stiffness of the haptic device. We incorporate the concept of decision-making models into the haptic assistance scenario, proposing two well-suited models that are integrated as the main negotiation strategy that forms part of an adaptive haptic-assistance scheme. The proposed assistance is tested in a partially structured environment.

Conditions

Haptic assistance design is a task-specific process, and may vary from task to task. We used a generic path-following task (the nominal task) with obstacles (un-modeled 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 the human-robot interaction. This generic scenario is the underlying principal of a series of real-life teleoperation applications (e.g. [Aarno et al., 2005, Boessenkool et al., 2013]), and also other scenarios of physical human-robot interaction (consider e.g. a mobility assistant [Yu et al., 2003], human-robot collaborative manipulation [Li et al., 2015, Mörtl et al., 2012], rehabilitation [Veras et al., 2008]) or training with a haptic aid [Li et al., 2009].

Previously published studies have focused on the evaluation of the natural adaptation of some proposed assistance compared with no-assistance or constant assistance schemes, and subjective questionnaires on human perception have been applied in order to measure the enhancement. However, there is no general agreement about how the final, *natural* assistance mechanism must be, hence this work proposes comparing the decision-making based assistance with a human-like assistance which consist of a human-expert coupled haptically with the tested subject.

Outline

This document is organized as follows: the first chapter presents briefly the related work for telerobotic systems, haptic assistances and gives an overview of concepts on dynamic decision making; chapter two establishes theoretical control concepts which support the varying admittance control law proposed; chapter three describes the components of the proposed assistance and introduces two assistance mechanisms (in this chapter we present the assistance dynamics when it interacts with a human in a path tracking task); each mechanism is based on a decision-making model borrowed from cognitive studies; chapter four establishes the main hypotheses and experimental methods; chapter five reports the experimental results and discussion; finally, chapter six gives some conclusions and future research topics.

2. Related work: haptic assistance in teleoperation

Telerobotic systems connect humans and robots in order to reproduce operator actions at a distance; these systems need to develop technology that deals with the inclusion of a human operator in the control loop of a remote robot [Ferre et al., 2007].

In robotics, autonomy is defined in terms of a system's ability to function effectively without human intervention [Dias et al., 2008]. As human operators may feel fatigue, irritability or cannot keep an optimal task execution, a better approach may be the design of an autonomous agent. However, in complex tasks, e.g. unknown and unstructured environments, the system may become too complex and needs human intervention.

When including humans into a complex automation system such as telerobotic systems, a dilemma rises as posed by Sheridan and Roseborough [Sheridan, 1992],

1. In a complex control system the controlled process cannot be fully and explicitly modeled, nor can the objective function.
2. In order to make simpler the control system, the human may be included into the control loop to supervise some functions
3. Since the human makes mistakes, it is evident that s/he can be assisted by a computer-based decision aid, so one must provide such aid.
4. To design the decision aid and evaluate the human operator's use of it, a relatively complete process model and objective function must be used as a norm. Step 4 is in conflict with step 1

Remark [Sheridan, 1992]**:** *If such a relatively complete process model and objective function were available, then why not use these in place of the human operator to provide an automatic decision maker, thus leaving the human out the control loop?*

Roseborough concluded that *"In any system requiring a human operator, the objective validity of a specific decision aid can never be established"*. On the other hand, Sheridan [Sheridan, 1992] proposed that the human decision maker is necessary in situations that are not explicitly modelable and no valid decision aid can be built to provide useful information when it is needed. Thus, it makes sense to provide decision aid for those situations where an explicit model function is available and thus the human can be benefited for these cases.

In this thesis, 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.

To provide such decision aid, an efficient communication between the operator and the robot is required. Efficient communication provides the chance to create a cooperatively plan for task execution and a negotiation policy to resolve conflicts. The decision aid may use audio, visual and/or haptic cues to communicate its states—and intentions—to the human operator.

In this thesis, we explore a novel idea borrowed from decision-making studies to assist the human operator via the haptic channel; the main goal is improving performance in teleoperation tasks without restricting the actions of the human operator.

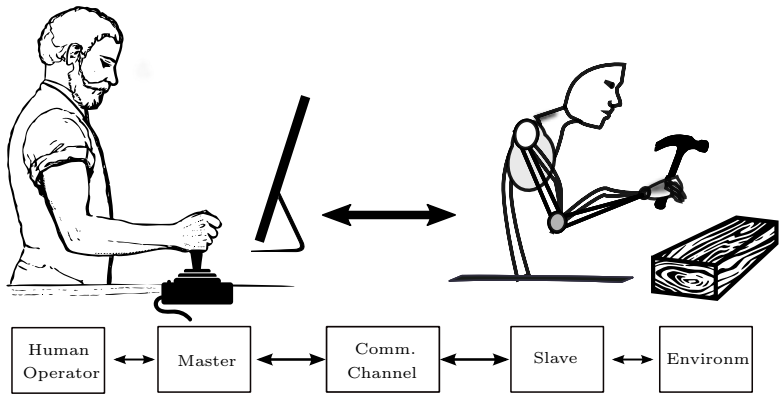


Figure 2.1.: Components of telerobotic systems. A telerobotic system consists of the human operator, the master device (haptic), a communication channel, the slave and the environment.

This chapter introduces three fundamental areas addressed in this thesis: (i) telerobotic schemes; (ii) haptic assistances; and (iii) decision making models. The first part surveys the telerobotic systems where a simple model is presented and the conditions to achieve an ideal telerobotic system are summarized. Next, haptic assistance techniques between human and robots is introduced. Finally, decision making models are presented.

2.1. Teleoperation

Telerobotics means literally "robot at a distance". The term *tele* is derived from the greek-root that means distant. The remote robot that physically performs the task and interacts with the remote environment is called *slave*, Fig. 2.1. The human may command the slave through a unilateral device that sends the commands to the robot, or a bilateral device that sends commands to the slave and at the same time receives information from the environment e.g. force; this device is called the *master*. In the human-machine interaction area, devices like keyboards, mouse and screens are unilateral devices because they act as input devices (e.g. keyboard and mouse) or output devices (e.g. displays, screens). On the other hand, haptic devices are bilateral devices because they act as both input and output devices. The operator transmits positions and/or velocities and at the same time receives forces—from a virtual or real world.

The unilateral and bilateral terms refer also to the type of the teleoperation system. A bilateral telerobotic system acts as an interface through which the operator sends commands and receives information from the remote environment. The operator acts in real-time on the remote environment and receives information from the interaction between the slave robot and its environment. In a unilateral scheme the information travel uniquely from the input device (master) to the slave, no kinesthetic feedback forces are available. On the other hand, in a bilateral scheme, the control information flowing bidirectionally between master and slave, and generally there is reflection of kinesthetic forces.

Besides reproducing forces exerted by the slave on the environment, the feedback may communicate other useful information e.g. by vibrotactitle cues, or additional forces that guides task execution.

In teleoperation, different control schemes are defined according to the type and number of information exchanged between master and slave, e.g. position-position, position-force or force-force; the two variables means two-channels indicating the flowing direction of the variable from master to slave and vice versa. A multichannel scheme transmits more than two variables e.g.

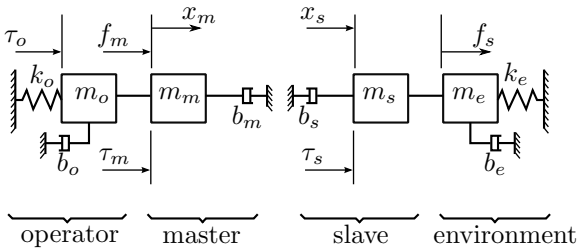


Figure 2.2.: A telerobotic system model composed by passive mechanical elements

velocity and force flowing in both directions, which uses a four-channel scheme.

2.1.1. A telerobotic system model

A telerobotic system consists of the human operator, the master device (haptic), a communication channel, the slave and the environment (Fig. 2.1). A common model of telerobotic systems incorporates mechanical systems [Yokokohji and Yoshikawa, 1994, Torres, 2007]. Modelling in Fig. 2.2 by k_o, b_o and m_o, which are the stiffness, damping and inertia of the human arm respectively and τ_o is the force generated by the muscles of the human. The master is also modeled as a second order system where m_m is the inertial mass and b_m is the damping factor of the master (viscous friction); f_m is the force applied by the operator to master and τ_m is the force of the master drive (actuator). The dynamics of the master device are modeled by,

$$\tau_m + f_m = m_m \ddot{x}_m + b_m \dot{x}_m \tag{2.1}$$

Considering no-displacement between the operator and the master [Yokokohji and Yoshikawa, 1994] i.e. the operator is attached rigidly to the master, the dynamic of the force exerted by the human operator is given by,

$$\tau_o - f_m = m_o \ddot{x}_m + b_o \dot{x}_m + k_o x_m \tag{2.2}$$

The human perception delays and the intrinsic bandwidth of the human motor control system may be included by using the McRuer model (crossover model) [Feth et al., 2009a], which is given by,

$$G_h = \frac{e^{-\tau_d s}}{1 + \tau_h s}[k_c(1 + \tau_c s)] \tag{2.3}$$

where, τ_d is the time delay due to the perception and τ_h is the lag due to the limited bandwidth of the human motor control system; k_c and τ_c are parameters of the control action [Feth et al., 2009a, McRuer and Jex, 1967].

The human-operator adapts k_c automatically (naturally) to achieve a stable interaction with the environment [McRuer and Jex, 1967]. τ_o does not depend on the state variables of the master [Yokokohji and Yoshikawa, 1994] therefore its common considering that τ_o does not present values that deem the system unstable.

Operator dynamics (G_o) may be modelled by a parameter-varying system which varies to interact with different environments [Yokokohji and Yoshikawa, 1994]. However, to simplify the analysis we considered G_o as a constant-parameter system.

The input for a teleoperation system simulation may be provided by the minimum jerk model [Flash and Hogans, 1985]. It consists of a fifth-degree polynomial that describes free movements of the human arm in a plane.

Similarly, the slave and the environment are described as second-order systems where m_s and b_s are the inertial mass and the damping factor of the slave, respectively; m_e, b_e and k_e are the inertia,

the damping (viscous friction) and the stiffness of the environment, respectively. The dynamics of the slave are given by,

$$\tau_s - f_s = m_s\ddot{x}_s + b_s\dot{x}_s \tag{2.4}$$

and the dynamic of the environment (considering that the slave is attached rigidly to the environment) is given by,

$$f_s = m_e\ddot{x}_s + b_e\dot{x}_s + k_e x_s \tag{2.5}$$

2.1.2. Ideal properties of telerobotic systems

Ideally, a telerobotic system should permit the operator to perform the task as if s/he conducted it directly as if no teleoperation system is involved.

Besides stability, which is the fundamental requirement for every control system, telepresence, or transparency, is one of the key factors that enhances performance of telerobotic systems. Telepresence means that the information available in the remote environment is displayed to the operator in a natural manner, which implies a feeling of presence at the remote site. A good degree of telepresence dictates the feasibility of the required manipulation task [Ferre et al., 2007, Hashtrudi-Zaad and Salcudean, 2001].

Telepresence involves the human senses and the natural synchronization of the different perception channels. Thus haptics is one of the elements involved in the concept of telepresence, but not the only one. We focus our descriptions on the haptic channel, which is created by the robotic hardware and its control system. The master–slave system becomes the medium through which the user interacts with the remote environment and, ideally, s/he is fooled into forgetting about the medium itself. If this is achieved, we say that the telerobotic system is **transparent** [Niemeyer et al., 2008].

To achieve transparency in the haptic channel three objectives for designing a teleoperation system were proposed in [Yokokohji and Yoshikawa, 1994, Lawrence, 1993],

1. Positions x_m and x_s, due to the operator's input τ_o, must be equal regardless of the dynamics of the environment.
2. Forces f_m and f_s, due to the operator's input τ_o, must be equal regardless of dynamics of the environment.
3. Both positions and forces, due to the operator input τ_o, must be equal regardless of the dynamics of the environment.

The impedance and the two port model concepts describe these three conditions in a compact representation [Hannaford, 1989, Lawrence, 1993, Yokokohji and Yoshikawa, 1994, Naerum and Hannaford, 2009]. Impedance relates velocity (position) and force, variables which are key to achieve transparency. The two-port model, on the other hand, describes the interconnection between master and slave.

A general control law

A general control law for 1 DoF telerobotic system (Fig. 2.3) is given in [Yokokohji and Yoshikawa, 1994],

$$\tau_m = \mathbf{P}_m\mathbf{x}_m + Q_m f_m - [\hat{\mathbf{R}}_m\mathbf{x}_s + \hat{S}_m f_s] \tag{2.6a}$$

$$\tau_s = \mathbf{R}_s\mathbf{x}_s + S_s f_s - [\hat{\mathbf{P}}_s\mathbf{x}_m + \hat{Q}_s f_m] \tag{2.6b}$$

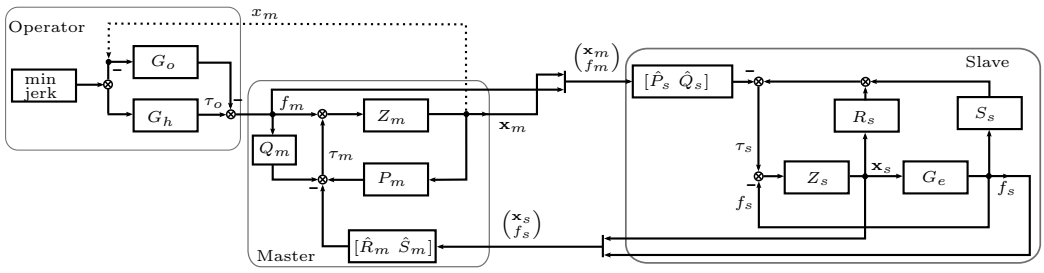


Figure 2.3.: A block diagram for a general control of telerobotic systems

where, $\mathbf{x}_*$ is a vector of the control variables given by,

$$\mathbf{x}_* = [x_*\ \dot{x}_*\ \ddot{x}_*]^T \text{ with } * \in [m\ \ s]$$

Q_* and S_* are scalars (in this 1 DoF example), and $\mathbf{P}_*$ and $\mathbf{R}_*$ are vectors of the form,

$$\mathbf{P}_*, \mathbf{R}_* \rightarrow [c_1^*\ c_2^*\ c_3^*]$$

The "hat" $(\hat{\cdot})$ indicates that the variable could include a time delay or scaling variables to make consistent operations (in case of asymmetrical configuration of master and slave devices).

The general control law (2.6) may represent common control architectures such as position-position (considering e.g. k_2, k_3 in $\hat{\mathbf{P}}_s, \hat{\mathbf{R}}_m$ and Q_s, S_m as zero) in which the master sends its position to the slave and the slave feedback its position to the master; or position-force control in which the master sends its position to the slave and the slave feedbacks the force exerted into its environment to the master.

Two-port network representation

To achieve a transparent teleoperation system a two-port network analysis is proposed. This framework is taken from the area of electrical circuit analysis (cf. e.g. [Nilsson and Riedel, 2004]). Consider a two-port network like the one shown in Fig. 2.4. In this representation, the source T_{op} corresponds to the force exerted by the operator; z_o is the operator dynamics (without the McRuer model); z_e is the load that represents the impedance of the remote environment and the two-port network is the teleoperation interface (master-communication-slave).

The current in the electrical domain (i) corresponds to the velocity in the mechanical domain $(\dot{x})$ i.e. for purposes of notation $\dot{x} \triangleq i$[1], and the voltage (v) corresponds with the force (f) i.e $f \triangleq v$.

From the expressions of the master (2.1), slave (2.4) and the general control law (2.6) (the correspondence between mechanical and electrical domains), we obtain,

$$T_m + F_m = (m_m s + b_m)\dot{X}_m \triangleq Z_m \dot{X}_m \tag{2.7a}$$

$$T_s - F_s = (m_s s + b_s)\dot{X}_s \triangleq Z_s \dot{X}_s \tag{2.7b}$$

where F_* is the force, $\dot{X}_*$ the velocity and Z_* the impedances for master and slave $(* \in [m, s])$. The upper case letter represents the Laplace transform of the corresponding variable.

Since $\mathbf{x}_* = \dot{X}_*$ in (2.6) (Fig. 2.3), then P and R take the form of,

$$P_*, R_* \rightarrow \frac{c_1^*}{s} + c_2^* + c_3^* s \text{ with } * \in \{m\ ,\ s\}$$

[1]In the analysis of electrical circuits, the current i_s i.e. $\dot{X}_s$ has positive direction, opposite to that shown in Fig. 2.4. Therefore, to take the results available from the network theory the negative sign $-\dot{X}_s$ is assumed

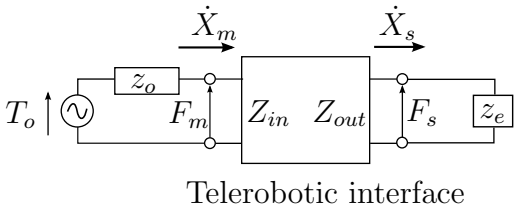


Figure 2.4.: Two-port network model. The human and remote environment connection.

The general control law (2.6) in matrix notation is given by,

$$\begin{aligned} T_m &\triangleq [P_m \quad Q_m] \begin{bmatrix} \dot{X}_m \\ F_m \end{bmatrix} - [R_m \quad S_m] \begin{bmatrix} \dot{X} \\ F_s \end{bmatrix} \\ T_s &\triangleq [P_s \quad Q_s] \begin{bmatrix} \dot{X} \\ F_m \end{bmatrix} - [R_s \quad S_s] \begin{bmatrix} \dot{X} \\ F_s \end{bmatrix} \end{aligned} \tag{2.8}$$

The impedance matrix ($\mathbf{Z}$) relates forces (voltages) and velocities (current) on each side of the two-port network and is given by [Nilsson and Riedel, 2004],

$$\begin{bmatrix} F_m \\ F_s \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} \dot{X}_m \\ -\dot{X}_s \end{bmatrix} \tag{2.9}$$

where the elements z_{ij} are given by [Yokokohji and Yoshikawa, 1994],

$$\begin{aligned} z_{11} &= \frac{(1+S_s)(Z_m - P_m) + S_m P_s}{(1+S_s)(1+Q_m) - S_m Q_s} &\triangleq \frac{N_{11}}{D_Z} \\ z_{12} &= \frac{-(1+S_s)R_m + S_m(Z_s - R_s)}{(1+S_s)(1+Q_m) - S_m Q_s} &\triangleq \frac{N_{12}}{D_Z} \\ z_{21} &= \frac{(1+Q_m)P_s + Q_s(Z_m - P_m)}{(1+S_s)(1+Q_m) - S_m Q_s} &\triangleq \frac{N_{21}}{D_Z} \\ z_{22} &= \frac{(1+Q_m)(Z_s + R_s) - Q_s R_m}{(1+S_s)(1+Q_m) - S_m Q_s} &\triangleq \frac{N_{22}}{D_Z} \end{aligned} \tag{2.10}$$

Different relationships between the variables involved in the two-port model can be derived e.g. given the impedance matrix we may obtain the admittance, hybrid or chain matrices, among other operators [Nilsson and Riedel, 2004].

The matrix "chain" ($\mathbf{K}$) represents the relationship between force and velocity on the slave side with their peers in the master side as follows,

$$\begin{bmatrix} F_m \\ \dot{X}_m \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} F_s \\ -\dot{X}_s \end{bmatrix}$$

where the elements of $\mathbf{K}$-matrix in terms of the elements of the impedance matrix are given by,

$$\mathbf{K} = \frac{1}{z_{21}} \begin{bmatrix} z_{11} & |\mathbf{Z}| \\ 1 & z_{22} \end{bmatrix} \tag{2.11}$$

The determinant of the impedance matrix is [Yokokohji and Yoshikawa, 1994],

$$|\mathbf{Z}| = \frac{(Z_m - P_m)(Z_s + R_s) + P_s R_m}{(1+S_s)(1+Q_m) - S_m Q_s} \triangleq \frac{D_Y}{D_Z}$$

The general control law (2.8) describes an ideal response (kinesthetic coupling) if its chain matrix representation corresponds to the identity [Yokokohji and Yoshikawa, 1994]. This matrix

represents the three conditions (cf. page 7) to obtain an ideal kinesthetic coupling. Thus, for a transparent telerobotic system its chain matrix must fulfill that,

$$\mathbf{K}_{\text{transparent}} \triangleq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

On the other hand, the hybrid matrix (**H**) [Lawrence, 1993,Hannaford et al., 1991,Niemeyer and Slotine, 1991,Anderson and Spong, 1989b,Naerum and Hannaford, 2009], describes the interaction between the force on the master side and velocities on the slave side in terms of the master velocity and the force exerted by the slave on its environment.

$$\begin{bmatrix} F_m \\ -\dot{X}_s \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} \dot{X}_m \\ F_s \end{bmatrix}$$

An ideal transparent teleoperation system has a hybrid matrix associated to the matrix [Naerum and Hannaford, 2009],

$$\mathbf{H}_{\text{transparent}} \triangleq \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

An ideal bilateral scheme

The general control law (2.8) that reproduces a transparent coupling must fulfil,

$$\mathbf{K}_{\text{transparent}} \Leftrightarrow \begin{bmatrix} \frac{N_{11}}{N_{21}} & \frac{D_Y}{N_{21}} \\ \frac{D_Z}{N_{21}} & \frac{N_{22}}{N_{21}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.12}$$

Elements, k_{11} and k_{22}, of the chain matrix are the transmission coefficients of force and velocity, respectively; elements k_{12} and k_{21} are known as the transmission impedance and admittance, respectively. Note that tracking velocity between master and slave is achieved if $D_Z = 0$ and $N_{22} = N_{21}$, while tracking forces is achieved if $D_Y = 0$ and $N_{11} = N_{21}$.

By inspection of (2.7), (2.8), (2.10) and (2.12) (cf. [Yokokohji and Yoshikawa, 1994]), $D_Y = 0$ only if $c_1^{m,s} = 0$, $c_{2,3}^{m,s} \neq 0$ and $\hat{c}_{2,3}^{m,s} \neq 0$ in P_m and R_s; the mass and damper of master and slave have to be known to include it into the control terms P_m and R_s. This means that the control law must include the acceleration term, but the acceleration measure may be difficult in practice. Besides, it requires that master and slave impedances are known, but in practice equal impedance cannot be achieved. If there is a difference between these values the system behaves as a negative mass.

Besides, in order to $D_Y = 0$, either $P_s = 0$ or $R_m = 0$. Then the following gains should be zero,

$$\hat{c}_1^m = \hat{c}_2^m = \hat{c}_3^m = 0 \quad \text{or} \quad \hat{c}_1^s = \hat{c}_2^s = \hat{c}_3^s = 0$$

This last expression reflects the idea that the system sends position variables $(x, \dot{x}, \ddot{x})$ in a direction and force variables in the other.

The control scheme presented here is a 4 channel model, as two variables are transmitted from master to slave ($\dot{x}$ and f) and the same number of variables in the opposite direction. A transparent teleoperation system with two channels, in which a channel transports velocity and the other force, was proposed in [Naerum and Hannaford, 2009]. Nonetheless it is necessary to know exactly the parameters of master and slave, which leads again to implementation issues.

In order to achieve $N_{11} = N_{21}$, which is other requiriment to achieve perfect tracking force behavior, restrictions in control gains values for local and fedforward forces have to be introduced. Similarly, in order to achieve perfect velocity tracking $D_Z = 0$ condition imposes restrictions in control gains values for local and fedforward forces. Besides, $N_{22} = N_{21}$ requires to know exactly mechanical parameters for master and slave in order to achieve perfect position tracking.

Transparent telerobotic interface, in practice, cannot be implemented. Nonetheless in theory, transparent control laws can be derived (cf. e.g. [Naerum and Hannaford, 2009, Lawrence, 1993, Yokokohji and Yoshikawa, 1994]). To cover this implementation issue an additional impedance is included, in order to modify the impedance of master or/and slave. For example, when the task is mainly composed of movements in free-space (without contact with the environment), a low impedance that reduces the physical effort of the human can be considered. On the other hand, when tasks require contact with the environment, the additional impedance may reproduce the dynamic of the remote environment. Thus the user perceives as if s/he is directly manipulating the remote environment. This additional impedance does not guarantee an ideal scheme, but in practice task performance is improved and facilitates the system implementation (at least if the parameters of the environment are known and constant).

Although including this impedance solves some implementation issues, in fact it brings new challenges such as determining the optimal impedance or adaptive mechanism that permits transferring the human skill to a remote site. Toward this end, several control strategies have been proposed, which may be classified based on the degree of authority (autonomy) given to the human on task execution.

2.1.3. Bilateral, supervision and shared control strategies

In teleoperation literature, three approaches are proposed to reflect the human skill at a remote site. The first consists in reproducing directly the actions executed by the operator at the remote site. The second provides the teleoperation system with a mechanism to perform the task autonomously while the human acts as a supervisor of the task. For the first approach the already presented bilateral paradigm is proposed and for the second a supervisor control scheme. The third approach refers to shared control, in which a combination of the above two approaches is proposed.

When the operator performs certain actions with full authority and others with assistance schemes, the control scheme is known as a shared control system; this scheme combines the best characteristics of bilateral control approaches and supervisor schemes. The human-operator executes actions directly and autonomously and the teleoperation system performs other actions to improve task performance. In this context the haptic assistance emerges to involve kinesthetic and tactile information to improve human-robot interaction.

Although operators have the ability to abstract new information, to learn and to make decisions in unexpected situations, when the operators perform the task via telerobotic interfaces, their skills are not transferred ideally to the remote site. Mainly because of technical restrictions in the telerobotic interface, such as the visual feedback which introduces perception error, the properties of the remote environment are generally unknown; also, there may exist a delay in the information that are exchanged between the operator and the remote site. Another restriction is the unnatural interaction that may exist between humans and robots due to the reduced number of degrees of freedom in robotic devices.

The shared control scheme solves the drawbacks and improves task performance by including information about the task, the operator and the remote environment simultaneously [Passenberg and Buss, 2010]. When this information is include in the control loop, the teleoperation system is known as an augmented teleoperation system [Unterhinninghofen, 2009]. On the other hand, if no additional information in the control scheme is included, the system is called a classic scheme [Hokayem and Spong, 2006].

Although adding extra information can reduce the realism of the perception, this only occurs if it is compared with a perfect, nonexistent, teleoperation system. However, in real systems with uncertainties (signals and model), the teleoperator may perform other actions different from those that operator requires to improve performance. In this case, altering the human actions improves the execution of the task, therefore the perception of realism is also improved [Unterhinninghofen, 2009, Boessenkool et al., 2011].

(a) Bimanual mobile haptic interface [Peer and Buss, 2008a]

(b) Two VISHARD10 haptic interfaces [Ueberle et al., 2004]

(c) Phantom Omni [3D-Systems,]

(d) Sigma 7 [Force-dimension,]

Figure 2.5.: Examples of force haptic devices

2.2. Haptics

Haptics is defined as the science that studies the sense of touch [Furht, 2006]; this term comes from the greek-root *haptikós*, that means to grasp. Haptic within the human-machine context introduces the sense of touch and force, that is tactile perception (from the latin, *tangere* that means to touch), which refers to the touching of surfaces. The broader definition of haptic includes the concept of kinesthesia (from the greek, *kinesis* that means movement and *aesthesis* that means perception), i.e. the sensing of movement in the body.

Haptics enhances the sensation of "presence" and provides information which cannot be described completely with only visual or audio feedback [Furht, 2006] such as stiffness and texture of objects. Therefore, haptics enables the user to manipulate objects in a environment—virtual or real—in a natural and effective manner.

Haptic interfaces cover a wide variety of devices. Their classification is based on the feedback component transmitted to the user [Ferre et al., 2007]. The predominant feedback component criterion permits classifying haptic devices into two categories: kinesthetic predominant devices—also known as force haptic devices, Fig. 2.5—and tactile predominant devices (tactile devices). This thesis will concentrate on kinesthetic devices that take advantage of force feedback to assist the human when performing a remote task.

Haptics has been explored in several contexts such as 3D modeling and animation, geophysical analysis, dentistry training, virtual museums, assembly planning, mine design, surgical simulation, design evaluation, control of scientific instruments, and robotic simulation. An overview of the most common types of haptic interfaces can be found in [Ferre et al., 2007].

2.2.1. Haptic Assistance

Haptic assistance explicitly refers to the use of haptic cues to assist the human operator. Although these cues may refer to vibrotactile information, in this thesis it corresponds to forces generated artificially, which guide the user when executing the task. In other words, this forces facilitate or restrict human movements in certain areas of the task space.

Haptic assistance is commonly implemented as virtual fixtures (see detail in Sec. 3.3.1). Virtual fixtures impose constrains which are overlaid on the physical environment in the workspace of the robotic device through the control system [Mihelj and Podobnik, 2012]; thus the user feels as if a real mechanical constraint exists. Two types of artificial forces can be generated in virtual fixtures: active and passive. In passive fixtures the force of the operator is scaled (augmented or reduced). Therefore, if the human does not exert any force on the haptic interface, then the robot keeps immobile. On the other hand, active fixtures consist on additional forces that are applied towards/against certain area of the workspace; thus the robot moves even if the human does not

exert any force [Passenberg et al., 2011].

The assistance design involves determining the instances when the human and/or the autonomous agent should intervene. In human-robot interaction argot this agent is called the *assistance*, authority or arbitration and it is implemented as shared control (also known as traded control).

The type of autonomous system mentioned so far is covered by a wide range of devices, from computer games controllers [Oguz et al., 2012, Enes and Book, 2010] to driver support systems for cars e.g. to assist steering in the vehicle or assist gas-pedal systems [Abbink et al., 2012]. Other autonomous systems that benefit from shared control are: mobile robots commanded at a distance [Desai and Yanco, 2005, Bruemmer et al., 2005, Shen et al., 2004, Medina et al., 2012], teleoperation interfaces [Passenberg and Buss, 2010, Boessenkool et al., 2011, Evrard and Kheddar, 2009b], haptic devices for surgery training [Nudehi et al., 2005, Powell and O'Malley, 2011, Kragic et al., 2005], anthropomorphic robots that interact directly with humans to transport an object jointly [Mörtl et al., 2012, Evrard and Kheddar, 2009a, Kronander and Billard, 2014] and mobility aids for the elderly [Yu et al., 2003].

Haptic studies commonly present human-machine interaction in structured environments, in which following a known path is usually the task at hand; this approach may include an on-line estimation of the path or the human intention [Aarno et al., 2005, Li and Okamura, 2003, Bettini et al., 2004, Veras et al., 2008, Weber et al., 2009], but in general the environment is assumed to remain unchanged. This generic scenario underlies a series of real-life teleoperation applications (e.g. [Aarno et al., 2005, Boessenkool et al., 2013]), but also other scenarios such as physical human-robot interaction (consider e.g. a mobility assistant [Yu et al., 2003], human-robot collaborative manipulation [Li et al., 2015, Mörtl et al., 2012], rehabilitation [Veras et al., 2008]) or training with haptic aids [Li et al., 2009] may benefit from such assistance systems.

Adapting the assistance level to unmodelled situations or uncertainty when tracking a path is of great importance in real life applications, but this experimental setup has not been widely considered (with the exception of [Passenberg et al., 2013]). In teleoperation this situations arise even in basic tasks, hence the importance of covering this drawback.

High-performance haptic interaction can be achieved by means of a constant assistance level when the improvement of just one objective is required [Passenberg et al., 2013] e.g. just improve task performance. Incorporating a constant assistance may improve task performance, but this is only true if the human operator agrees with the proposed assistance mechanism. Otherwise, the user fights against the proposed assistance and the comfort decreases, degrading performance. A constant assistance level (e.g. [Nudehi et al., 2005, Desai and Yanco, 2005].) may not function well to improve multiple design objectives e.g. in partially structured environments, due to the fact that the assistance does not adapt to non-programmed situations; the same assistance level is provided for known and unknown situations. For a high assistance level, the user may perceive lack of task control; while for a low assistance level the user has more authority on task execution making it easy to overrule the programmed assistance, but perceives absence of assistance when it is required.

An optimal assistance for unstructured or partial structured environments requires a mechanism to solve this trade-off between performance and control effort [Groten, 2011, Passenberg et al., 2013, Kucukyilmaz et al., 2013, Dragan and Srinivasa, 2013, Abbink et al., 2012]. In order to solve this trade-off, varying the assistance level is proposed as the main mechanism to improve multiple design objectives. Incorporating such varying strategies in a human-robot interaction system may lead to a more human-like interaction [Groten, 2011]. Such varying mechanisms are presented when two humans interact through haptic feedback. The humans naturally adopt a specialization of roles which change during task execution [Reed and Peshkin, 2008].

The assistance level may be established by trial and error or by means of pilot experiments that include 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 [Marayong et al., 2002], or by solving a multi-objective problem e.g task performance, comfort and effort [Passenberg

et al., 2013].

On-line adaptation of the assistance level has been proposed in [Kucukyilmaz et al., 2013] by heuristics implemented as state machines or in [Abbink et al., 2012, Evrard and Kheddar, 2009b, Dragan and Srinivasa, 2013, Enes and Book, 2010] by the construction of a scheduling function. This function, commonly noted as α, estimates the assistance level based on metrics of task performance [Passenberg et al., 2013], interaction with the user [Medina et al., 2012, Dragan and Srinivasa, 2013] or any other task information [Evrard and Kheddar, 2009b, Desai and Yanco, 2005, Kragic et al., 2005].

The benefits of having varying assistance schemes over constant schemes (or non-assisted configurations) are explored in [Boessenkool et al., 2013, Abbink et al., 2012, Passenberg et al., 2013, Mörtl et al., 2012]. Varying assistance improves task and/or comfort—effort—performances compared to constant assistance, but the users adapt better to constant assistance [Passenberg et al., 2013, Mörtl et al., 2012]. This might suggest that the user cannot decode the proposed varying assistance mechanism while in constant assistance the user deduces easily the mechanism. Recently, concepts borrowed from game theory have been proposed in [Kucukyilmaz et al., 2013, Li et al., 2015] to give the robot a more human-like reaction; this approaches will be explored in next Sec. 2.3.

2.3. Decision making

Decision making is the process of choice making amongst different alternatives. This process is complex and comprehends several research areas like behavioral sciences (psychology, cognitive science), social sciences (economics, political), neuroscience, biology, game theory, computer science and it is recently explored in robotics. Decision-making can be classified into normative theories which investigates what the decision maker should do, descriptive theories that explain what the decision maker really does and prescriptive theories which combines normative and descriptive approaches [Coskunoglu and Weber, 1989].

Normative strategies refer to the obtention of an optimal policy to achieve a known goal; this policy is closely related to decision theory and utility theory. Although normative theories may describe experimental data that match human-behavior, they do not explain its relation to the underlying decision-making mechanism [Egelman et al., 1998] and generally assume non-real human behaviors like rationality—humans always maximize their own payoff. For instance, decision theory is a set of analytical techniques that support the decision-maker's choice such that it is optimal among a set of alternatives; game theory, on the other hand, supports the decision-maker when the preferred outcomes are clearly defined and known but the optimal outcome depends on the strategy adopted by other decision-makers. Both theories consider rationality and utility paradigms.

Descriptive theories, on the other hand, explain real human behaviors e.g. suboptimal strategies, irrational behaviors or intransivity. For instance, behavioral decision-making studies when the user seeks non-optimal outcomes; regret theory considers that the consequences of a decision are not independent of each other, and it is possible that choices are in contradiction with the transivity assumption, i.e. if the operator prefers choice option A to B and B to C, then s/he must prefer A to C.

Prescriptive theories combine normative and descriptive theories to study how humans make decisions taking into account real human-behaviors. Nowadays neuronal activity can be recorded and related with optimal-or-suboptimal strategies. As is pointed in [Bogacz et al., 2007], neural behavior reveals the decision-making mechanism that humans (and animals in general) use; this relatively new information exposes the evolutionary mechanism and therefore the optimal behavior which humans adopt under pressure and in trade-off situations.

We are interested in the *dynamic decision making* approach, which studies the interdependent decision-making process. These sequence of decisions occurs in an environment that changes over the time because of previous decisions made by the decision-maker or inherent changes in the environment dynamics. While common approaches provide a unique and conventional one-time

decision, in dynamic decision making the decision is more complex and changes with time, requiring more experienced users to control the changing environment.

Providing the robot with a decision-making mechanism enhances its autonomy as well as human-robot interaction. If the robot is able to adopt more human-like reactions, the human may understand better the robot's response and perceive the interaction as more natural. Although the decision making process is a very active research area, the incorporation of such process to provide more natural and autonomous human-robot interaction is on its first attempts.

2.3.1. Decision making and haptic assistances

Recently, concepts borrowed from game theory have been successfully applied to coordinate humans and robots movements via a haptic channel. When coordinated movements are required, a common cost function for task performance is introduced; then, an optimal control law based on *Nash's equilibrium* is derived to control position and force [Li et al., 2015]. When the human disagrees with the provided assistance, the control law is adapted to minimize the error between the force exerted by the user and the optimal force. Oguz et.al. [Oguz et al., 2012] proposed a scheme to solve collaborative, but also conflictive situations in a haptic interaction by three negotiation models—borrowed also from game theory: concessive, competitive and tit-for-tat. In *concessive* negotiation a computer-agent makes concessions for the benefit of the human; in the *competitive* model the computer-agent regards its interests more than those of the human; and in the *tit-for-tat* model the computer-agent gives a series of concessive steps until it notices an inadequate reaction in the human decisions, point at which the computer-agent adopts retaliation actions.

The human-robot interaction described so far are based on rationality and utility concepts of the decision-making models. In contrast, the *Drift-Diffusion (DD) model*—which rather than considering utility or rationality, seeks to describe a possible non-optimal utility result in a dynamic decision making scenario—is incorporated into a supervision scheme in which a human supervises a team of robots that exploits and explores an unstructured environment [Cao et al., 2008, Stewart et al., 2012]. Nevertheless, no-haptic feedback is included.

2.3.2. Dynamic decision making

Dynamic decision making refers to a sequence of interdependent decisions (decisions that are interrelated) in a temporal order. The decisions take place in an environment that changes autonomously or in response to the decision maker's actions [Woodruff et al., 2012, Fischer et al., 2015, Gonzalez, 2013, Gonzalez and Quesada, 2003].

The cognitive scientists construct dynamic decision-making models to determine when the humans adopt a particular strategy in the long term, achieving optimal decisions or exploring different ones in an attempt to maximize their rewards. These studies are of interest in economics, for example, to assist investors to make the best decision to maximize the income or rewards [Bodnaruk and Simonov, 2015]. Recently, a decision-making mechanisms to aid air traffic controllers has been proposed, because of the large amounts of information that they need to process affecting optimal decision-making. Hence, it is argued that an agent capable of making decisions may improve the performance of the human actions [Hill et al., 2005].

According to cognitive studies, the humans make a decision when sufficient evidence is accumulated favoring one alternative over the other [Egelman et al., 1998]. This mechanism fits neuronal and behavioral data in the simplest decision situation: choosing between two options, the so called *two-alternative force-choice task* (TAFCT). This basic task reveals important aspects of the decision-making mechanism e.g. it shows that humans adopt a sub-optimal strategy, called the *matching point*, at which both choices result in the same reward (payoff), Fig. 4.3. Only few subjects find the optimal sequence that achieves better performance.

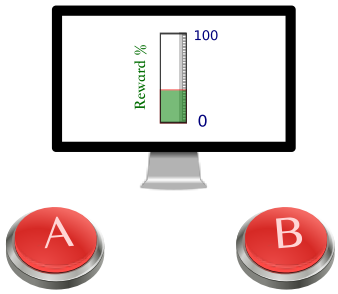


Figure 2.6.: Cognitive experiment, the two-alternative force choice task. In the TAFCT, the human sequentially chooses between two options A or B and s/he is awarded a reward for her/his choice; the reward is shown proportional to the height of the bar. The goal of the task (game) is that the operator accumulates at the end of the experiment the highest reward.

The TAFCT studies two experimental paradigms: the interrogation and the free-response paradigms. In the first case the lowest error rate is expected at a given decision time; in the latter the shortest reaction time is expected for a given error rate [Bogacz et al., 2006]. *The optimal behavior in the cognitive experiments refers to the model that reproduces the human behavior in these two paradigms rather than the model that optimizes the utility in the decision process.*

Two force alternative choice tasks

Montalgue et.al. [Montague and Berns, 2002] proposed a sequential economic game in which the human makes a series of decisions between two alternatives (A or B, Fig. 2.6). Every time the human makes a decision, s/he obtains a reward (performance measure). The goal of the game is to maximize the overall reward.

Although decision making is a complex stochastic process, choosing between two alternatives is a simplification of many real-life situations and its study is well accepted within the dynamic decision-making community because [Bogacz et al., 2006]:

- It represents real-life problems to which animals are challenged in their natural environment, e.g. move towards or away from an unknown stimulus. In these situations, the decision-making time and accuracy can mean life or death, so the decision mechanism has a strong evolutionary influence, therefore optimal.
- A large amount of data concerning human behavior has been generated from as early as since the nineteenth century, which has led to formal and well accepted formulation of dynamic models for the TAFCT.
- Neuroscientists can now monitor neural dynamics and evaluate their relationship to task performance. Neuronal and behavioral data (in the TAFCT experiment) converge to support formal models, such as the Drift-Diffusion Model (DD).

In decision-making studies, three considerations for modeling the TAFCT are proposed [Bogacz et al., 2006]:

- The evidence in favor of each alternative is accumulated over time (involves memory).
- The process is subject to random fluctuations; there exists nondeterministic elements on the environment and human variation cannot be known beforehand.
- A decision is made when enough evidence is accumulate in favor of one alternative over the other. According to many cognitive studies, considering the difference between the

accumulated evidence rather than accumulate evidence of each choice separately, gives better results.

Next, the costs associated to each alternative are modeled by *the reward structures* and then a brief description of the decision making models proposed in the dynamic decision making area is presented.

Reward structures

The two alternatives in choice situation correspond to two activities in which the human has to exploit available resources or find new ones. The cost of the resources is modeled by two functions whose outputs consist of a reward for each alternative; these two functions are known as a reward structures. The reward structure models a trade-off in the cost of exploiting the resources or exploring for new ones e.g. the human tries to increase the reward exploiting a resource, but over the time the resource decreases so the human has to change his/her strategy to explore for new ones, and again increase the reward.

The reward obtained by the human at each decision-making step is a deterministic function $r(t)$—unknown to the human—which is given by

$$r(t) = \begin{cases} r_A & \text{if } z(t) = \text{A} \\ r_B & \text{if } z(t) = \text{B} \end{cases} \tag{2.13}$$

where $z(t)$ is the decision made at time t ($z \in [\text{A B}]$), A and B are the two-alternatives and $r(t)$ the corresponding reward. The reward is calculated based on the last reward and the history of previous rewards; if the human chooses option A, the reward is calculated based on the reward function r_A, otherwise, the reward is calculated based on the reward function r_B. The dependence on past decisions—modeled by r_A and r_B—is relevant to solve real problems in which the human makes decisions when interacting with complex systems [Stewart et al., 2010] e.g. aircraft, air traffic control, human-robot cooperative manipulation, teleoperation.

Several reward functions (r_A and r_B) have been proposed in cognitive science (Fig. 4.3 shows four). Rather than representing specific real-world tasks, the different structures are intended in isolating and testing specific behaviors of human decision-making mechanisms [Woodruff et al., 2012]. For instance, in the converging Gaussian (Fig. 2.7c) the subject is pushed to the matching point, while in the diverging Gaussians (Fig. 2.7d) subjects are encouraged to explore. Both structures match the optimal strategy with the matching point [Stewart et al., 2012]. For the rising optimum structure (Fig. 2.7b), the optimal behavior is difficult to find for a specific test subject as there exists a local optimum ($\psi = 0$) and a global optimum ($\psi = 1$); even more the matching point is located near the local optimum. The matching shoulder structure (Fig. 2.7a) generalizes these behaviors [Nedic et al., 2008], where the matching point and the global optimum not necessarily coincide; therefore, different behaviors can be obtained by the linear case according to the chosen parameters.

To maximize the reward of the cognitive task, an optimal sequence of decisions has to be made. The optimal strategy is the one that maximizes the average reward function $g(\psi) = \psi r_A + (1-\psi) r_B$; the average is presented in Fig. 2.7 as the optimal label for four common reward structures.

The matching shoulder reward structure (Fig. 2.7-a) is given by

$$\begin{aligned} r_A &= \text{a}_1\psi + \text{a}_2 \\ r_B &= \text{b}_1\psi + \text{b}_2 \end{aligned} \tag{2.14}$$

where, in cognitive studies, r_A, r_B are calculated based on the *allocation of* "A", here represented by ψ. The allocation of "A" is the amount of choices made of a particular option (A) in the last N decision making situations, therefore, if $N = 20$ then $\psi = \#\text{A's}/20$. The parameters a_1, b_1 are the slopes of the lines in Fig. 2.7a and they represent the ratio of the change in the available resources.

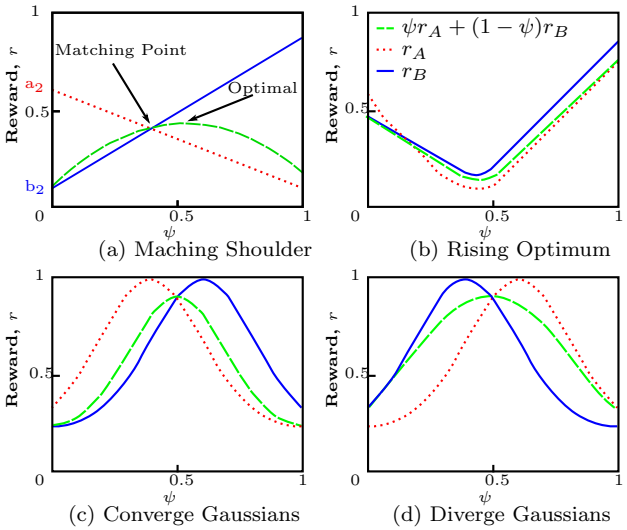


Figure 2.7.: Four reward structures or functions [Stewart et al., 2012]

The parameter a_2 is the maximum amount of resources (reward) available for the option "A" i.e. when the resources for "A" are not exploited, the amount is high, therefore, the reward is a_2. On the other hand, b_2 is the amount of resources exploited situations the decision was "B" in the previous N situations; in this case, b_2 represents the lowest reward due to successively choosing the resource "B", i.e. the resources for "B" are exploited repeatedly so the amount of choices "B" are low, representing a low reward of value b_2.

So far, we presented the economic game that cognitive scientists have proposed to study the human decision-making process in TAFCT. Cognitive scientist [Montague and Berns, 2002, Bogacz et al., 2007] have determined the decision-making mechanism through experiments and have used behavioral and neurobiological arguments to justify the likelihood of the mechanism proposed [Cao et al., 2010]. The human-mechanism consists on deciding when sufficient evidence is accumulated favoring one alternative over the other [Egelman et al., 1998]. Next, some models that represent this mechanism are presented.

2.3.3. Decision-making models

A simple rule such as the win-stay-lose-switch, WSLS (also known as γ-policy), reproduces the matching point behavior (which may lead to a suboptimal sequence of decisions). In this rule, the decision-maker changes its choice if the current reward decreases, and keeps his/her current choice otherwise; the WSLS rule in TAFCT may be interpreted as a finite state machine [Woodruff et al., 2012]. WSLS rule involves only the two most recent rewards, although reproduces the matching point behavior, it deviates from human mechanisms such as learning from past decisions based on accumulated experience. Some users use this memory-based strategy to find the optimal strategy that maximizes the average reward.

The dynamic decision-making is also embedded into the context of Markov decision processes [Meisel, 2011], in which the current state serves as the evidence for the decision. To cover the interdependence of the decisions, a Markov model with a predictive stage is proposed—called anticipatory behavior—to find an optimal trade-off between immediate contribution of the current decision and its impact on future contributions. The optimal decision is recursively calculated based on the expected value of the remaining decision process (in terms of state transition probabilities and Bellman equation) [Meisel, 2011]. Other approaches based on Markov processes model the

attention to a particular option at each moment, which reflects the decision maker's underlying subjective probability (or belief) in favor of a particular alternative; this can be seen as a deliberation process [Johnson and Busemeyer, 2006].

In the same line of thought, an artificial neural network has been used to represent this deliberation process in which a classic weighted additive utility model is used, but the attention weights are stochastic, representing the fluctuation in the evaluation of the possible consequences for each alternative; formally this can be seen as a Markov process [Busemeyer and Johnson, 2004].

Another artificial neural network relates attributes with choices in the so-called ECHO model [Tsuzuki and Guo, 2004], where the connection weight between an attribute node and a choice node is determined by the value of the choice with that attribute. This model predicts that as one option becomes dominant during deliberation, the activation of the attribute nodes favored the dominant alternative.

Stochastic differential equations have been used to represent a connectivist approach, differing in the use of one or more integrators to accumulate the evidence. The connectivist term refers to a graphical model that emulates the neuronal behavior observed in the decision-making process. For instance, in the race (inhibition-free) model the accumulators for each of the two alternatives integrate evidence independently. The Mutual inhibition model (also known as competing accumulator or LCA-Leaky competing accumulator) is composed by two leaky accumulators; each accumulator is connected with a mutual inhibitory connection. The feedforward inhibition model presents integrators with no leak, and they receive inhibition from "crossed" inputs rather than inhibiting each other. Decisions are made whenever the activity of either accumulator reaches a threshold.

On the other hand, two models that accumulate the difference (instead of treating them indepently) of the evidence supporting each choice are the Drift-diffusion (DD) models and the Ornstein-Uhlenbeck (OU) model; the difference is that in OU the evidence also depends on the current value of the difference and it can accelerate or decelerate toward the threshold, similar to the ECHO model, in which a dominant choice is selected. Conditions that allow to describe the DD-model as a Markov process are presented in [Stewart et al., 2010].

Amongst the previous models, the DD model is popular because of its simplicity and success in fitting behavioral and neuronal data in several human studies (see [Bogacz et al., 2006] and references therein). Besides, the drift diffusion model is the optimal decision-maker in the two aforementioned experimental situations: interrogation and free response. Actually the OU, Mutual inhibition and feedforward inhibition achieve their optimal behavior for parameters that reduced then to a DD model (the race model cannot reach optimal behavior) [Bogacz et al., 2006]. Here optimality refer to the model-behavior which fitting human behavior on free-response and interrogation paradigm, instead of achieve an optimal response in the sense of classical approaches as utility-based response.

We incorporate two models into the haptic assistance scenario, namely the WSLS rule and the Drift-diffusion model. A more detailed description for these models is presented next.

The Win Stay Lose Switch Model (WSLS)

The WSLS model takes into account the previous reward in order to make decisions. Therefore, when the current reward is greater than the previous reward, the model keeps choosing the same option, but for a lower or equal reward the model changes its choice.

The decision rule is given by [Cao et al., 2008]

$$z(t+1) = \begin{cases} z(t) & \text{if } r(t) > r(t-1) \\ \sim z(t) & \text{otherwise} \end{cases} \tag{2.15}$$

where $z(t)$ is the decision at time t ($z(t) \in$ [A B]), $r(t)$ the reward at time t and the symbol $\sim$ denotes the "not" operator; i.e. if $z(t) =$ A (resp. $z(t) =$ B), then $\sim z(t) =$ B (resp. $\sim z(t) =$ A).

The first case in (2.15) is called an exploitation situation while the second is an exploration of resources. The exploration option improves the reward when seeking new resources, and full understanding of this behaviour is in itself an area of research in behavioural studies of animals and humans.

The WSLS model is also called the $\gamma-$policy which is a special case of (2.15), when the first condition in (2.15) considers the greater or equal than the previous one, i.e. the exploration occurs just when the actual reward is strictly reduced compared to the previous obtained [Woodruff et al., 2012, Vu and Morgansen, 2008].

A convergence analysis is presented in [Vu and Morgansen, 2008] from the perspective of dynamical systems and in which the sequences of decision-making are modelled as a state machine. The model converges to a percentage of a choice, the so-called matching point (Fig. 2.7).

Drift-Diffusion (DD) Model

The DD model describes the human behavior in a sequential decision-making process. The human accumulates past rewards and makes predictions of future rewards in order to making decisions.

The neuronal behaviour in animals suggests that the change in the level of dopamine[2] represents the prediction error and the amount of future reward stimuli [Egelman et al., 1998]. The DD is a formal model based on physiological experimental data of the neuronal behavior of dopamine in primates and foraging bees. This computational interpretation suggests that from a behavioral perspective, increasing the level of dopamine is associated with a state of "better than expected" and decreasing the level with a "worse than expected" state.

Montague et.al. [Montague and Berns, 2002] performs cognitive experiments showing that the DD model adjusts and predicts better the data obtained in humans experiments.

The DD model is a reinforcement learning model, in which evidence in favor of one choice over the other is integrated until a predetermined threshold is reached [Nedic et al., 2008].

The simplest version of DD model is described by a stochastic differential equation [Nedic et al., 2008] that describes a process used widely to model perceptual decision making [Bogacz et al., 2006]. This version is presented in the context of multi-choice decision making dynamics, therefore in one-dimensional situation the evidence in favor of a choice of interest is accumulated by [Cao et al., 2010]

$$dz = \vartheta dt + \sigma dW \ ; \ z(0) \tag{2.16}$$

where z represents the accumulated evidence in favor of one choice, ϑ denotes the drift rate, this is the signal intensity of the stimulus acting on z and σdW the increments drawn from a Wiener process (noise) with standard deviation σ which is the diffusion rate representing the effect of the noise.

Considering a TAFCT with A and B choices, the drift rate ϑ is determinated by subject's anticipated rewards for a decision A or B, denoted as w_A and w_B, these are the prediction or reward expected [Egelman et al., 1998, Bogacz et al., 2007, Cao et al., 2010].

In TAFCT the accumulated evidence (z) is computed by the difference of two choices. The choice (A or B) is made when $z(t)$ first crosses any of the predetermined thresholds $\pm z_{th}$.

The decision rule called soft-max has the same form as that predicted by the Drift-Diffusion Model in TAFCT [Stewart et al., 2012]. The soft-max model predicts the probability of choosing an option A by [Bogacz et al., 2007]

$$p_A(t+1) = \frac{1}{1+\exp^{-\mu(w_A - w_B)}} \tag{2.17}$$

[2]Dopamine is a neurotransmitter. A neurotransmitter is a substance produced by a nerve cell capable of altering the operation of another cell in a short or prolonged time.

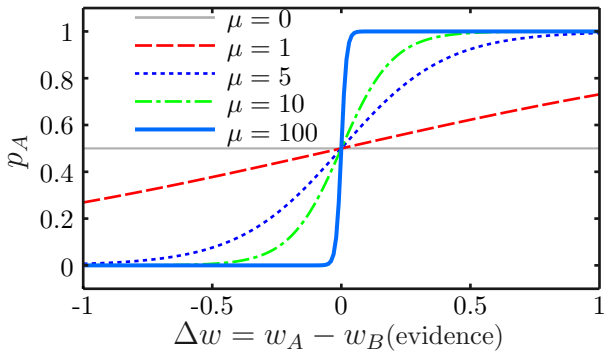


Figure 2.8.: Certainty-parameter (μ) in the probability of choosing an option (A)

The slope of the sigmoidal function is determined by the parameter μ (Fig 2.8). Large μ implies more certainty in the decision making, which can be interpreted as less tendency to explore. If μ tends to infinity, p_A becomes deterministic (see $\mu = 100$ in Fig. 2.8).

A relationship between the models (2.17) and (2.16) is presented in [Bogacz et al., 2006, Cao et al., 2010], in which $\mu = \frac{z_{th}}{\vartheta}$ and $w_A - w_B = \frac{\vartheta}{\sigma}$.

The weights $w_A(t)$ and $w_B(t)$ are the expected rewards for choices A and B, respectively. The difference between w_A and w_B is the evidence in favor of choosing A, referred as Δw.

The expected reward, proposed in the learning theory area [Bogacz et al., 2007], is given by

$$w_z \leftarrow w_z + \lambda(r - w_z) \tag{2.18}$$

where $z \in [\mathrm{A}, \mathrm{B}]$ and represents the decision just made, $r(t)$ is the actual reward and $\lambda \in [0, 1]$ is the learning rate; a larger λ means less memory.

An error between the actual reward (r) and the expected one (w_z) is calculated (2.18) and weighted by the learning rate (λ). Other expression proposed to accumulate the evidence is given by [Stewart et al., 2012]

$$w_z(t+1) = (1 - \lambda)w_z(t) + \lambda r(t) \tag{2.19}$$

where the evidence is accumulated in favor of the decision z just made.

Both equations (2.19) and (2.18) represent the idea of accumulating the weight of the choice just made.

3. Haptic assistance control scheme

Our primary aim is to propose a varying haptic assistance; hence, in this chapter we stablish theoretical conditions to vary the impedance of a common teleoperation scheme.

The stability of teleoperation systems has been extensively studied. Where linear control systems theory ([Fite et al., 2004, Lawrence, 1993, Mobasser and Hashtrudi-Zaad, 2008, Shahdi and Sirouspour, 2009], [Christiansson et al., 2006, Daniel and McAree, 1998], [Gil et al., 2004, Hulin et al., 2008, Azorin et al., 2001]), Lyapunov methods [Salcudean and Zhu, 2000, Strassberg et al., 1993, Sirouspour and Shahdi, 2006] and passivity theory [Anderson and Spong, 1989a, Hashtrudi-Zaad and Salcudean, 2001] have been applied. Gain scheduling in robotics has been also tackled in many different ways, including look-up tables [Huang and Chang, 2004], LPV and polytopic control [Patton and Klinkhieo, 2010] and adaptive control strategies [Kelly et al., 1989].

The application of this gain scheduling techniques has been generally focussed on enhancing performance of nonlinear systems. The direct application of such techniques to address haptic assistance in teleoperation context is problematic since the goal is to assist a human operator in order to achieve task performance, while still retaining interaction properties that give full authority to the operator over the system. Even more, previous work on haptic assistances has reported that an optimal strategy with multi-criteria performance measurements may result unnatural to the user as it is not clear for the feedback which of the multiple measures requires improvement [Passenberg et al., 2013]. Besides, task performance decreases if the human and the proposed assistance disagree; evenmore, this increase human effort [Kucukyilmaz et al., 2013, Mörtl et al., 2012] or even the task may not be accomplished. This fact complicates the implementation of a gain scheduling strategy as a closed form solution, bringing to the foreground the need for new scheduling techniques in order to incorporate the intricacies of human-machine interaction. In this direction we propose the use of **D**ecision **M**aking **M**odels (DMMs) that are found naturally within the context of cognitive science and human-human interaction, and uses a model of such negotiation as a scheduling technique which varies the amount of impedance imposed by the assistance.

The fact that we are using a mechanism naturally adopted by humans, somehow supports later conclusions where operators prefer the use of a DMM based-assistance strategy over no assistance or a human-expert assistance strategies.

This chapter presents theoretical concepts needed to implement the gain scheduling control scheme in a real teleoperation system. For this purpose we will address the simplest control problem where we make the following assumptions:

A1. The robot manipulators at the master and slave sides have an exact models.

A2. The interaction between the slave and the environment is strictly passive (SP).

A3. The haptic interaction between the human and the master device is stable.

A3. The communication channel has no time delays.

Assumption A1 is needed in order to achieve an exact feedback linearization (see Appendix A for a short review on computed torque technique taken from [Spong et al., 2005, Craig, 2005]); this compensation present both at the master and slave sides, thus the robot plant (G_R considered in Fig. 3.1) is assumed to be linear from now on. This (linear) plant is later augmented with a PD (internal loop) controller in order to obtain nominal position tracking. If we want to imprint some type of "reference" dynamics to the system, e.g. if we want to vary the system's apparent stiffness, an additional block must be introduced. The additional filter (Y) implements the desired dynamics

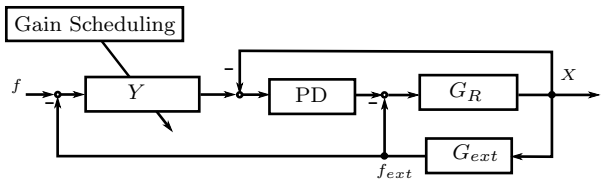


Figure 3.1.: Position-based admittance control, varying admittance (Y), G_R haptic device and G_{ext} operator dynamics. The admittance (Y) is a varying parameter and implements the assistance level in the haptic task.

and outputs a reference position (see Fig. 3.1). This position-based admittance controller is used as the main mechanism to vary the assistance level perceived by the human operator.

Stability of the overall teleoperation architecture may be concluded from the following two steps:

1. The master side is stable and strictly passive 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 (A3).
2. The second step uses the previous fact to conclude that under assumptions A1 and A4, the teleoperation system is stable since we have the feedback interconnection of two strictly passive components.

Notice that these theoretical results have a strong impact in the design of haptic assistance since it states that, under assumptions A1-A4, the designer may vary the parameters of the impedance without restriction and focus on the interaction and task performances requirements.

Before presenting the proposed varying admittance scheme, we define passivity and affine varying systems, concepts by which theoretical conditions for the system stability can be derived. Sec. 3.2 outline basic concepts to conclude stability of the studied variable impedance teleoperation controller under common assumptions made in the telerobotics field.

3.1. Passive systems

In telerobotics, both the master and the slave interact with its environment. The master interacts with the human and the slave with the physical environment. The passivity theory gives a formal framework to study the telerobotic system as the connection of passive systems (this is an active area of research in robotics see e.g. [Secchi et al., 2007, Ortega et al., 1998]).

For instance, the dynamics of the operator's arm is adaptable and time varing [Hogan, 1989], often modelled approximately by a linear time-invariant (LTI) mass-damper-spring system, about its operating point, in series with an external force source [Raju et al., 1989, Hannaford, 1989]. The environment can be highly varying and nonlinear (e.g. intermittent contact, free motion to hard contact and vice-versa). It is often assumed that the slave makes contact with the environment, which is modelled by an LTI mass, damper and spring system about its operating point [Hannaford, 1989, Lawrence, 1993]. The master and slave are also modelled as LTI mass (and damper) systems, but nonlinear models have been also used [Strassberg et al., 1993, Chopra and Spong, 2004, Zhu and Salcudean, 2000].

Passivity is a powerfull tool for analysis of complex systems such as telerobotics; this framework applies to linear and nonlinear systems. The components of the teleoperation systems can be represented as power-port elements, e.g. the communication channel with time delay may result in a passive network [Niemeyer and Slotine, 1991, Anderson and Spong, 1989a]; the uncertainties such as most interaction objects are within the passive class and, since the operator's contact with any strictly passive object is stable, operators are assumed to present a passive property [Colgate and

Hogan, 1988]. Several research to give a formal treatment of the coupled stability has been done (see e.g. [Anderson and Spong, 1989b, Colgate, 1993, Adams and Hannaford, 1999, Hashtrudi-Zaad and Salcudean, 2001, Haddadi and Hashtrudi-Zaad, 2010]). Next we define passivity and some properties of interconnections of passive systems (taken from [Márquez, 2003, Khalil, 2001]). This is done in order to present a self-contained document, nonetheless its application in our proposed scheme is straightforward.

Definition 1 *[Khalil, 2001] The system*

$$\dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}, \mathbf{u}) \tag{3.1a}$$
$$\mathbf{y} = \boldsymbol{h}(\mathbf{x}, \mathbf{u}) \tag{3.1b}$$

is passive if there is a continuously differentiable positive semidefinite function $V(\mathbf{x})$, called storage function, such that,

$$\mathbf{u}^T\mathbf{y} \geq \dot{V} = \frac{\partial V}{\partial \mathbf{x}}\boldsymbol{f}(\mathbf{x}, \mathbf{u}), \quad \forall(\mathbf{x}, \mathbf{u}) \tag{3.2}$$

Besides, the system (3.1) is said to be

- lossless if $\mathbf{u}^T\mathbf{y} = \dot{V}$
- input strictly passive if $\mathbf{u}^T\mathbf{y} \geq \dot{V} + \mathbf{u}^T\boldsymbol{\varrho}(\mathbf{u})$ for some function $\boldsymbol{\varrho}$ such that $\mathbf{u}^T\boldsymbol{\varrho}(\mathbf{u}) > 0, \forall \mathbf{u} \neq 0$
- output strictly passive if $\mathbf{u}^T\mathbf{y} \geq \dot{V} + \mathbf{y}^T\boldsymbol{\gamma}(\mathbf{y})$ for some function $\boldsymbol{\gamma}$ such that $\mathbf{y}^T\boldsymbol{\gamma}(\mathbf{y}) > 0, \ \forall \mathbf{y} \neq 0$
- strictly passive if $\mathbf{u}^T\mathbf{y} \geq \dot{V} + \mathbf{u}^T\boldsymbol{\zeta}(\mathbf{x})$ for some positive definite function $\boldsymbol{\zeta}$

Example 1. Mass, spring and damper system

Consider a one-dimensional mechanical system with a mass, a spring and a damper. The equation of motion is given by,

$$m\ddot{x}(t) + b\dot{x}(t) + kx(t) = u(t), \quad x(0) = 0, \dot{x}(0) = \dot{x}_0 \tag{3.3}$$

where m is the mass, d is the damper constant, k is the spring stiffness, x is the position of the mass and u is the force acting on the mass.

The system (3.3) is defined by its state space realization,

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -\frac{k}{m} & -\frac{b}{m}\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + \begin{bmatrix}0\\ \frac{1}{m}\end{bmatrix}u(t), \tag{3.4a}$$
$$y(t) = \begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} \tag{3.4b}$$

where the state variables are $x_1 = x$ and $\dot{x}_1 = x_2$, which correspond to position and velocity, respectively.

Intuitively we may expect that the rate of energy supplied by the system ($u^T y$), has a lower bound determined by the total energy of the system represented as the storage function for the system in (3.3). Thus, the total energy of the system—the sum of the kinetic and potential energy—is a candidate for V (which is a semi-definite positive function), as follows,

$$V(x_1, x_2) = \frac{1}{2}kx_1^2 + \frac{1}{2}m{x_2}^2 \tag{3.5}$$

The time derivative of V along the trajectories of (3.4a) is given by,

$$\dot{V} = \frac{\partial V}{\partial \mathbf{x}}\dot{\mathbf{x}} \tag{3.6a}$$

$$= \begin{bmatrix} kx_1 & mx_2 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ -\frac{k}{m}x_1 - \frac{b}{m}x_2 + \frac{u}{m} \end{bmatrix} \tag{3.6b}$$

$$= -bx_2^2 + x_2 u \tag{3.6c}$$

Since the output is the velocity x_2 ($y = x_2$), the inequality (3.2) holds for (3.3). Therefore, the system is passive. Even more, the mass, spring, damper system is output strictly passive, since there always exists a $\gamma < b$ such that $u^T y > \dot{V} + \gamma y^T y$.

Commonly the nonlinear model of a robot can be linearising and decoupled via the so-called computed torque control technique (Appendix A). The linear model obtained makes it possible to design a simple servo control, usually a PD controller. Under assumption A1 the resulting feedback system corresponds to a mechanical system, which is analogue to the system in (3.3).

The stability of the closed loop system with a PD controller can be established using energy arguments [Brogliato et al., 2010, Bao and L. Lee, 2007]. The proportional action corresponds to the spring force, and the derivative action corresponds to the damper force. As in the mechanical system the derivative action will dissipate the virtual energy that is initially stored in the system, and the unforced system will converge to the equilibrium origin.

Example 2. n-link robot

Consider the nonlinear dynamical system of an n-link robot given by [Haddad and Chellaboina, 2008]

$$\mathbf{u} = \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) \tag{3.7a}$$

$$\mathbf{y} = \dot{\mathbf{q}} \tag{3.7b}$$

where $\mathbf{q}(\mathbf{0}) = \mathbf{q}_0, \dot{\mathbf{q}}(\mathbf{0}) = \dot{\mathbf{q}}_0, t \geq 0$, $\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}} \in \mathbb{R}^n$ represent generalized position, velocity, and acceleration coordinates, respectively, $\mathbf{u} \in \mathbb{R}^n$ is a force input, $\mathbf{y} = \dot{\mathbf{q}} \in \mathbb{R}^n$ is a velocity measurement, $\mathbf{M}(\mathbf{q})$ is a positive-definite inertia matrix function for all $\mathbf{q} \in \mathbb{R}^n$, $\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$ is an $n \times n$ matrix function accounting for centrifugal and Coriolis forces and has the property that $\dot{\mathbf{M}}(\mathbf{q}) - 2\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$ is skew-symmetric[1] for all $\mathbf{q}, \dot{\mathbf{q}} \in \mathbb{R}^n$, and $\mathbf{g}(\mathbf{q})$ is an n-dimensional vector accounting for gravity forces and is given by $\mathbf{g}(\mathbf{q}) = [\frac{\partial V(\mathbf{q})}{\partial \mathbf{q}}]^T$, where $V(\mathbf{q})$ is the system potential. We assume $V(\mathbf{0}) = 0$,$V(\mathbf{q})$ is positive definite, and $\mathbf{g}(\mathbf{q}) = \mathbf{0}$ has an isolated root at $\mathbf{q} = \mathbf{0}$.

To show that 3.7 is lossless [Haddad and Chellaboina, 2008], consider the energy storage function $V_s(\mathbf{q}, \dot{\mathbf{q}}) = \dot{\mathbf{q}}^\mathbf{T}\mathbf{M}(\mathbf{q})\dot{\mathbf{q}} + 2\mathbf{V}(\mathbf{q})$. Now,

$$\begin{aligned} \dot{V}_s(\mathbf{q}, \dot{\mathbf{q}}) &= 2\dot{\mathbf{q}}^\mathbf{T}\mathbf{M}(\mathbf{q})\dot{\mathbf{q}} + \dot{\mathbf{q}}^\mathbf{T}\dot{\mathbf{M}}(\mathbf{q})\dot{\mathbf{q}} + 2\frac{\partial V(\mathbf{q})}{\partial \mathbf{q}}\dot{\mathbf{q}} \\ &= 2\dot{\mathbf{q}}^\mathbf{T}[\mathbf{u} - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} - \mathbf{g}(\mathbf{q})] + \dot{\mathbf{q}}^\mathbf{T}\dot{\mathbf{M}}(\mathbf{q})\dot{\mathbf{q}} + 2\mathbf{g}^\mathbf{T}(\mathbf{q})\dot{\mathbf{q}} \\ &= 2\dot{\mathbf{q}}^T\mathbf{u} + \dot{\mathbf{q}}^T\left(\dot{\mathbf{M}}(\mathbf{q}) - 2\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\right)\dot{\mathbf{q}} \\ &= 2\dot{\mathbf{q}}^\mathbf{T}\mathbf{u} \end{aligned} \tag{3.8}$$

thus, the system is lossless ($\dot{\mathbf{q}} = \mathbf{y}$). Alternatively, with $\mathbf{u} = -\mathbf{K_p}\dot{\mathbf{q}} + \mathbf{v}$, where $\mathbf{K_p}$ is a positive-definite matrix, it follows that

$$\dot{V}_s(\mathbf{q}, \dot{\mathbf{q}}) = 2\mathbf{y}^\mathbf{T}\mathbf{v} - 2\mathbf{y}^\mathbf{T}\mathbf{K_p}\mathbf{y}, \tag{3.9}$$

[1] given any vector $\mathbf{x}$ a skew-symmetric matrix ($\mathbf{M}$) holds that, $\mathbf{x}^T\mathbf{M}\mathbf{x} = 0$

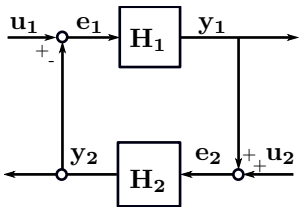


Figure 3.2.: Feedback of two systems

and hence, the input-output map from $\mathbf{v}$ to $\mathbf{y}$ is output strict passive. In order to prove asymptotic stability we use Lyapunov arguments.

Note that if $\mathbf{v} = \mathbf{0}$ then the Lyapunov candidate V_s has time derivative $\dot{V}_s(\mathbf{q}, \dot{\mathbf{q}}) = -2\dot{\mathbf{q}}^{\mathbf{T}}\mathbf{K_p}\dot{\mathbf{q}}$; to establish stability we apply Lasalle's theorem. First define the set $\mathcal{M} = \{\dot{\mathbf{q}}, \mathbf{q} \in \mathbb{R}^n : \dot{V}_s = 0\}$, and observe that $\mathcal{M}$ is the set of all states where $\dot{\mathbf{q}} = \mathbf{0}$. Now assuming $\mathbf{u} = \mathbf{0}$, and observing that a state of minimal energy (i.e. $(\dot{\mathbf{q}}, \mathbf{q}) \in \mathcal{M}$) is reached for $\mathbf{q} = \dot{\mathbf{q}} = \mathbf{0}$, the $\mathbf{g}$ term must be zero, which in turn is only true if $\mathbf{q}(\mathbf{0}) = \mathbf{0}$. Hence, $\mathcal{M} = (0, 0)$ (i.e. the origin) is asymptotically stable. Finally, we note that if $V(\mathbf{q})$ is radially unbounded, then global asymptotic stability is ensured.

3.1.1. Interconnection of passive systems

Consider the feedback interconnection of the systems $\mathbf{H_1}$ and $\mathbf{H_2}$ as in Fig. 3.2,

Theorem 1 *[Khalil, 2001] If H_1 and H_2 are passive then feedback connection is also passive*

Proof.
Let V_1 and V_2 be storage functions for systems $\mathbf{H_1}$ and $\mathbf{H_2}$ respectively. This implies that, if the systems are passive, then

$$\mathbf{e}_i^T \mathbf{y}_i \geq \dot{V}_i, \text{ for } i \in \{1, 2\} \tag{3.10}$$

From the feedback connection, we see that,

$$\begin{aligned} \mathbf{e_1^T y_1} + \mathbf{e_2^T y_2} &= (\mathbf{u_1} - \mathbf{y_2})^{\mathbf{T}} \mathbf{y_1} + (\mathbf{u_2} + \mathbf{y_1})^{\mathbf{T}} \mathbf{y_2} \\ &= \mathbf{u_1^T y_1} + \mathbf{u_2^T y_2} \end{aligned} \tag{3.11}$$

which implies that

$$\mathbf{u^T y} = \mathbf{u_1^T y_1} + \mathbf{u_2^T y_2} \geq \dot{V}_1 + \dot{V}_2 \tag{3.12}$$

Therefore, the storage function for the interconnected system is

$$V(\mathbf{x}) = V_1(\mathbf{x}) + V_2(\mathbf{x}) \tag{3.13}$$

We can conclude the feedback system is passive.

A similar procedure may be achieved to conclude that parallel connection of passive systems is passive.

Theorem 2 *[Khalil, 2001] Consider a feedback connection of two dynamical systems. When $u = 0$, the origin of the closed-loop system is asymptotically stable if each feedback component is either, strictly passive, or output strictly passive (OSP) and zero-state observable (ZSO). Furthermore, if the storage function for each component is radially unbounded, the origin is globally asymptotically stable.*

Prof [Khalil, 2001]
H_1 is SP; H_2 is OSP & SZO

$$\mathbf{e_1^T y_1} = \dot{V}_2 + \zeta_1(\mathbf{x_1}), \qquad \zeta(\mathbf{x_1}) > 0, \forall \mathbf{x_1} \neq 0 \tag{3.14}$$
$$\mathbf{e_2^T y_2} = \dot{V}_2 + \mathbf{y_2}^T \boldsymbol{\gamma}(\mathbf{y_2}), \qquad \mathbf{y_2^T}\boldsymbol{\gamma}(\mathbf{y_2}) > 0, \forall \mathbf{y_2} \neq \mathbf{0} \tag{3.15}$$

Thus,

$$\begin{aligned}
\mathbf{e_1^T y_1} + \mathbf{e_2^T y_2} &= (\mathbf{u_1} - \mathbf{y_2})^\mathbf{T}\mathbf{y_1} + (\mathbf{u_2} + \mathbf{y_1})^\mathbf{T}\mathbf{y_2} = \mathbf{u_1^T y_1} + \mathbf{u_2^T y_2} \\
V(\mathbf{x}) &= V_1(\mathbf{x_1}) + V_2(\mathbf{x_2}) \\
\mathbf{u} = \mathbf{0} &\Rightarrow \dot{V} \leq \zeta_1(\mathbf{x_1}) - \mathbf{y_2}^T\boldsymbol{\gamma_2}(\mathbf{y_2}) \\
\dot{V} = 0 &\Rightarrow \mathbf{x_1} = \mathbf{0} \text{ and } \mathbf{y_2} = \mathbf{0} \\
\mathbf{y_2}(t) \equiv \mathbf{0} &\Rightarrow \mathbf{e_1}(t) \equiv \mathbf{0} \ (\& \ \mathbf{x_1}(t) \equiv \mathbf{0}) \Rightarrow \mathbf{y_1}(t) \equiv \mathbf{0} \\
\mathbf{y_1}(t) \equiv \mathbf{0} &\Rightarrow \mathbf{e2}(t) \equiv \mathbf{0}
\end{aligned}$$

By ZSO of $\mathbf{H_2}$, $\mathbf{y_2} \equiv \mathbf{0} \Rightarrow \mathbf{x_2} \equiv \mathbf{0}$ and applying the invariance principle the connection is UAS.

3.2. Variable gain admittance control

We present some useful definitions.

Definition 2 *Convex set [Boyd and Vandenberghe, 2004]: A set S in a vector space over $\mathbb{R}^n$ is called a convex set if the line segment joining any pair of points of S lies entirely in S.*

Definition 3 *Convex hull [Boyd and Vandenberghe, 2004]: The convex hull of a set of n-dimensional points S in n is the intersection of all convex sets containing S. For N points $\mathbf{p}_1, \ldots, \mathbf{p}_N$, the convex hull C is then given by the expression,*

$$C = \sum_{j=1}^{N} \lambda_j \mathbf{p}_j : \lambda_j \geq 0 \quad \forall j \text{ and } \sum_{j=1}^{N} \lambda_j = 1.$$

Definition 4 *Convex polytope [Boyd and Vandenberghe, 2004]: A convex polytope may be defined as the convex hull (or convex envelope) of a finite set of points (which are always bounded)*

Definition 5 *Affine linear systems [Azhmyakov et al., 2011]: A system is affine in its control inputs (or control-affine), if a system over $\mathbb{R}$ takes the form,*

$$\dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}) + \boldsymbol{g}(\mathbf{x})\mathbf{u} \tag{3.16}$$

By definition all Linear Time Invariant (LTI) systems are affine systems.

The affine linear parameter varying system (aLPV) considered in this thesis is built considering a number of point-wise affine (linear) systems; each one describes the system locally in an operating point. If these models are combined in an appropriate way, that is by taking operating point dependent convex combinations of parameter values that belong to the different linear models, then an aLPV system will result. Consequently, the parameter values of a aLPV system vary within a polytope, and the vertices of this polytope are the parameter values that belong to the different linear models.

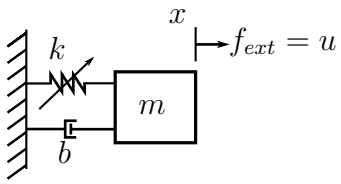


Figure 3.3.: Varying stiffness model for the rendered admittance

3.2.1. Affine linear parameter varying system

Since our formulation relays on the use of an admittance control schemes, we will use the mass-spring-damper system as an example throughout the section. Consider the system in Fig. 3.3. 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 given by,

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -\frac{k(t)}{m} & -\frac{b}{m}\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + \begin{bmatrix}0\\ \frac{1}{m}\end{bmatrix}u(t), \tag{3.17a}$$

$$y(t) = \begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}. \tag{3.17b}$$

Since our main objective is to vary the assistance level dynamically, first we must ensure that the time-varying control loop, as made explicit by the term $k(t)$ in the A-matrix in (3.17), is strictly passive for any bounded variation of $k(t)$.

First 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 [Kim and Ahn, 2011],

$$\dot{\mathbf{x}} = \sum_{i=1}^{N} \xi_i A_i, \qquad \sum_{i=1}^{N} \xi_i = 1, \tag{3.18}$$

In general, stability of a LTV system depends not only on the stability of the system for all (fixed) values of the parameter under consideration, but also on the dynamics of the parameter (i.e. the velocity of the variation must in general be small). Nonetheless, for a-LPV systems, stability of the autonomous system (3.18) may be guaranteed if there exists a Common Quadratic Lyapunov Function (CQLF) such that,

$$A_i^T P + P A_i < 0 \quad \forall i. \tag{3.19}$$

If this is the case, stability of (3.18) is guaranteed under *any* variation of the parameter [Boyd et al., 1994] (i.e. the velocity of $k(t)$ is not restricted). If $N = 2$ in (3.18), and the system is second-order, the existence of a CQLF is guaranteed under certain conditions as described in the next theorem.

Theorem 3 *[Lin and Antsaklis, 2009, Shorten et al., 2004]: Given two Hurwitz matrices A_1, $A_2 \in \mathbb{R}^{n\times n}$ such that* $\text{rank}(A_2 - A_1) = 1$*, 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.*

The rank condition on $(A2 - A1)$ means that only one parameters vary in one direction. Notice that in the haptic assistance scenario, the difference $(A_2 - A_1)$ is always rank $= 1$ by construction; and that eigenvalues for the system considered $(A_1 A_2)$ are,

$$\lambda_{1,2} = -\frac{k_1 m + k_2 m - b^2 \pm \sqrt{\eta}}{2m^2} \quad \text{with} \tag{3.20}$$

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

In order to guarantee stability of the gain scheduled system we require the eigenvalues of $(A_1 A_2)$ to have no negative real part. Thus, for the existence of a CQLF, sufficient conditions may be obtained, such that this is always true, i.e.

$$\eta < 0 \qquad k_1 m + k_2 m - b^2 < 0. \tag{3.21}$$

Finally, we may conclude that the interconnection of the robot and the aLPV-controller presented in Fig. 3.1 is stable if conditions (3.21) are satisfied.

3.3. Haptic assistance mechanism

If we assume that the low-level position controller has a fast and well damped dynamics, then a position-based admittance controller may be used to imprint via Y some reference dynamics to the master device (Fig. 3.1). Thus, the human perceives that the system is more or less stiff or damped. Varying either the stiffness (k) or damping (b) permit changing the systems dynamics, which can be translate into a more or less assistive control strategy; hereafter we will assume that only k is varied.

Last assumption is imposed because in our implementation we decided on purpose for a very generic path following with obstacles, instead of introducing a specific application-oriented task. The known path describes the nominal task and the obstacles represent unmodeled situations. In this generic task, the assistance is implemented as a guidance virtual fixture [Bettini et al., 2004], which varying the stiffness in the perpendicular direction of the desired path.

The haptic assistance proposed is implemented as a set of guidance virtual fixtures (GVF) [Bettini et al., 2004], which constrains the operator's motion in certain directions while allowing motion along the others directions. This scheme is widely used in the haptic assistances literature [Marayong and Okamura, 2004, Passenberg et al., 2013, Abbott and Okamura, 2006, Aarno et al., 2005, Li et al., 2007], therefore its application may be extended to a wide variety of scenarios.

3.3.1. Guidance virtual fixtures

A virtual fixture is an artificial, software-based, constraint that the operator feels as a mechanical or haptic constraint due to the fact that the robot's movements are limited.

Rosenberg [Rosenberg, 1993] defined virtual fixtures as

> "Abstract sensory information overlaid over the reflected sensory feedback from an environment with which the manipulator of the teleoperation or collaborative robotic system is in contact"

Virtual fixtures can be classified in two categories: guidance virtual fixtures (GVFs) and forbidden-region virtual fixtures [Abbott et al., 2005]. The former provides assistance to the user in moving the robot along a desired paths or surfaces, while the latter prevents motions into forbidden areas. Guidance assistances may be an active or passive scheme; in the active schemes a force is exerted towards and along the desired path [Pezzementi et al., 2007], such that if the user does not apply a force the robot still moves to accomplish the task; in the passive case, the force is scaled to guide the user to the desired path such that if the user does not apply any force the robot does not move.

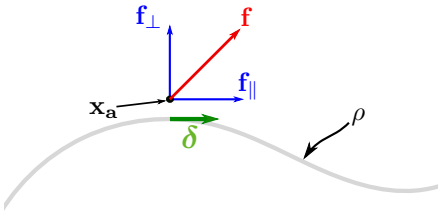


Figure 3.4.: Decomposition of the user force ($\mathbf{f}$) on the desired path ρ for the Guidance Virtual Fixtures scheme [Bettini et al., 2004]. $\boldsymbol{\delta}$ is the desired direction of the movement along the desired path ρ, $\mathbf{f}$ is the force exerted by the user, $\mathbf{f}_\perp$ and $\mathbf{f}_\parallel$ are the perpendicular and parallel components of the force and $\mathbf{x_a}$ is a point outside the path ρ.

The proposed set-up consist of an active guidance virtual fixture in the perpendicular direction of a desired path and a passive one for the parallell direction. For this purpose define a parametric path along the desired assistance trajectory, which serves us as a nominal task description, by a function ρ [Bettini et al., 2004]

$$\rho : [0\ 1] \rightarrow \mathbb{R}^3$$

It is assumed that ρ is differentiable, i.e.

$$t(s) = \frac{d\rho(s)}{ds}$$

where $s \in [0, 1]$ is the parameter. As s increases, a point in ρ moves on the path in a given direction ($\boldsymbol{\delta}$), Fig. 3.4.

Given a point $\mathbf{x_a}$ that does not belong to the path ρ, the closest point to the path is defined as $\hat{s}(\mathbf{x_a})$ by [Abbott and Okamura, 2006],

$$\|\rho(\hat{s}(\mathbf{x_a})) - \mathbf{x_a}\| = \min_{s\in[0\ \ 1]} \|\rho(s) - \mathbf{x_a}\|$$

The direction of the movement ($\boldsymbol{\delta}$) for each point $\mathbf{x_a} \in \mathbb{R}^3$ is defined by the tangent unit vector at the closest point on the curve, Fig. 3.4.

$$\boldsymbol{\delta}(\mathbf{x_a}) = \frac{t(\hat{s}(\mathbf{x_a}))}{\|t(\hat{s}(\mathbf{x_a}))\|}$$

Considering that the force exerted by the operator on the end effector of the master is available ($\mathbf{f}$), the aim is to create a virtual force that guides the operator to follow the nominal path ρ. For this purpose the projection operator is defined by [Marayong et al., 2003],

$$D \equiv \boldsymbol{\delta}(\boldsymbol{\delta}^T\boldsymbol{\delta})^{-1}\boldsymbol{\delta}^T$$

Note that $\mathbf{f}$ can be composed into two orthogonal components of force by,

$$\mathbf{f}_\parallel = D\ \mathbf{f} \tag{3.22a}$$
$$\mathbf{f}_\perp = \mathbf{f} - \mathbf{f}_\parallel \tag{3.22b}$$

where, $\mathbf{f}_\parallel$ is the force in the parallel direction to the nominal curve (tangent), $\mathbf{f}_\perp$ is the force exerted in the perpendicular direction to the curve, and $\boldsymbol{\delta}$ is the direction of the movement on the curve (a unit vector). These components of the force are related to the values of the admittance, one for the tangential direction ($Y_\parallel$) and another for the perpendicular direction ($Y_\perp$).

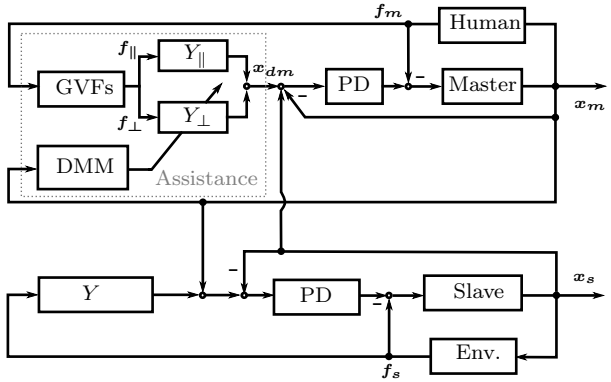


Figure 3.5.: Control scheme of the decision making based assistance, the assistance is composed by the decision making model (DMM) and the guidance virtual fixture (GVFs). The control scheme is a variable admittance (Y) and a PD controller

3.3.2. Teleoperation control scheme

A general admittance control law can be given by,

$$\dot{\mathbf{x}}_\parallel = Y_\parallel \mathbf{f}_\parallel \tag{3.23a}$$
$$\dot{\mathbf{x}}_\perp = Y_\perp \mathbf{f}_\perp \tag{3.23b}$$

In the generic task proposed here, the decision making model defines the suitable admittance on the perpendicular direction. Thus, the human is assisted to tracking the desired (nominal) path. For this purpose, the control scheme is composed by a variable admittance controller Y and a PD controller, Fig. 3.5. The varying admittance implements locally (at the master side) the assistance which consist of a guidance virtual fixture and a decision making model. The user interacts directly with the master device while the slave interacts with the remote environment. Both devices are sincronized via a position-based admittance control scheme [Peer and Buss, 2008b] and exchange position variables in the cartesian space.

The force measured at the tip of the master can be decomposed (3.22) into,

$$\mathbf{f}_m = \mathbf{f}_\parallel + \mathbf{f}_\perp, \tag{3.24}$$

where $\mathbf{f}_\parallel, \mathbf{f}_\perp$ 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, not force him to move along it, i.e.the user has full control authority of the task in the parallel direction of the desired path.

The admittance in the parallel direction ($Y_\parallel$) is given by

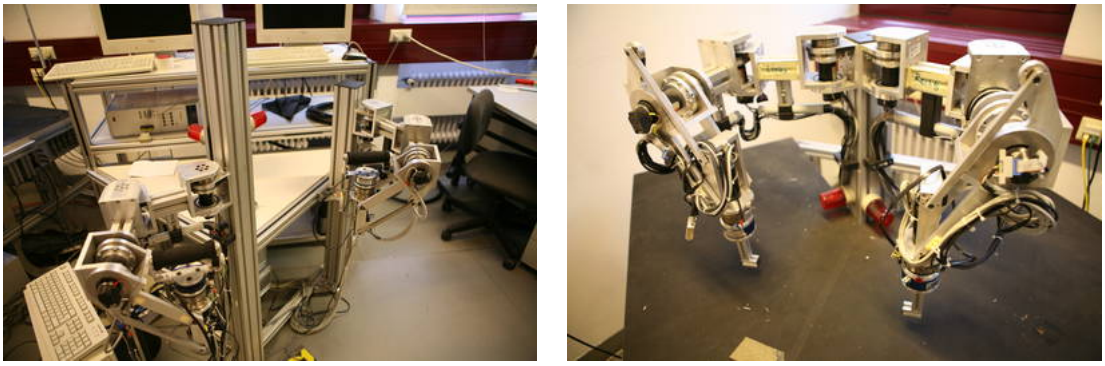
$$\mathbf{f}_\parallel = m_\parallel \ddot{\mathbf{x}}_\parallel + b_\parallel \dot{\mathbf{x}}_\parallel, \tag{3.25}$$

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

$$\mathbf{f}_\perp = m_\perp \ddot{\mathbf{x}}_\perp + b_\perp \dot{\mathbf{x}}_\perp + k_\perp \mathbf{x}_\perp, \tag{3.26}$$

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

$$\mathbf{x_{dm}} = \mathbf{x}_\parallel + \mathbf{x}_\perp. \tag{3.27}$$



(a) Master device — DeKiFeD (b) Slave device — DeKiToP

Figure 3.6.: Bimanual haptic telepresence system [Kron and Schmidt, 2005].

A passive GVFs is implemented in the perpendicular direction of the desired path, while an active GVFs is implemented in the perpendicular direction. Therefore, the human feels a force towards the desired path when moves out of the path, and the robot moves along the path only if the user exerts a force in the parallel direction. The active GVFs is constant during task execution.

Note that the assistance makes a decision about the stiffness value ($k_{\perp}$) that is suitable to improve the nominal task based on certain system variables. The decision making model (presented in the next chapter) adapts the assistance level, changing the stiffness of the master admittance in the perpendicular direction of the desired path (see $Y_{\perp}$ in Fig. 3.5).

3.4. Experimental hardware

The decision making model-based assistance was implemented on a real teleoperation system with haptic feedback devices (Fig. 3.6). The teleoperation system is a bimanual haptic telepresence system; it is composed by the master called DeKiFeD (Desktop Kinesthetic Feedback Device) and the slave called DeKiToP (Desktop Kinesthetic Teleoperator) [Kron et al., 1999, Kron and Schmidt, 2005]. The interface is located at Institute of Automatic Control Engineering at the Technical University of Munich (Germany); the manipulators are located in the same room and the communication between them is established by the Local Area Network of the institute. The system can be programmed using the Matlab-RTW with the RTAI-Target.

The teleoperator system consists of four admittance-type haptic feedback devices each one with four DoF. To validate the proposed assistance, only three DoFs were used. The kinematic and technical specifications of the teleoperator (Fig. 3.6b) are the same as the ones for the haptic input device (Fig. 3.6a). A six DoF force/torque-sensor (JR3) is mounted at the tip of the manipulators to measure interaction forces with the environment and the human operator respectively.

The bimanual teleoperator provides two arms in the master side and two in the slave side. In the haptic assistance test we use just one manipulator at the master's side that commands one manipulator at the slave's side. We take advantage of this bimanual teleoperator system to propose a multiple operator single robot (MOSR) scheme to evaluate a human-human interaction condition in order to compare the human expert assistance strategy with the decision making model-based assistance proposed (for details see the methods Chapter 5).

4. Decision making models on haptic assistances

Shared control haptic assistance improves the user's performance in telemanipulation tasks [Abbott and Okamura, 2006, Boessenkool et al., 2013, Unterhinninghofen, 2009]. Nevertheless, the design of haptic assistance policies for shared autonomy strategies is challenging because of the inherent difficulty in modeling unexpected situations e.g changes in the user intention, unstructured—or partial structured—environments or fault in the system.

According to cognitive science studies, humans make decisions based on previous rewards obtained on past decisions and estimate future rewards in order to maximize their total intake. This thesis takes inspiration from this paradigm to incorporate a similar behavior in our haptic assistance mechanism. The haptic assistance based on decision-making models makes a sequence of decisions on the suitable level of assistance needed to improve task and/or human perception. In short, the decision making model proposed accumulates evidence based on the current reward (r) and may include memory of previous ones, Fig. 4.1.

This chapter presents the components of the decision making model and its application to the haptic assistance context; we present the trade-off in haptic assistance as a TAFCT, its reward structure and how it influences the haptic assistance mechanism. Finally, we present two decision making models proposed initially in cognitive science as the dynamic (or policy) varying the assistance level.

4.1. Trade-off in haptic assistance

If the task proceeds according to plan (only nominal task occurs), without any unmodeled situations and the assistance is properly designed for this circumstance, the *agreement* between human and assistance about the plan to execute the task has to be high throughout the interaction. Nonetheless, if the task is too complex for the assistance, unexpected and/or unmodeled situations occur, this is not the case. When these situations occur, the assistance must improve the interaction without restricting the operators desire.

In such cases, the haptic assistance mechanism has to solve a trade-off, which consists in not providing unwanted assistance (over-assisting) and not failing to provide needed assistance (under-assisting) to the user [Abbink et al., 2012, Dragan and Srinivasa, 2013]. High assistance levels improve task performance if the human operator agrees with the assistance and no fault (or unexpected situations) is presented in the automation system. Otherwise, the human effort increases since s/he has to fight against the proposed assistance. On the other hand, for low assistance levels

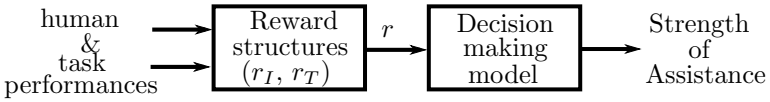


Figure 4.1.: The decision-making model (DMM) varies the assistance level (autonomy or authority) in the haptic task, based on the reward structures r_T and r_I, which are functions of human and task performance metrics.

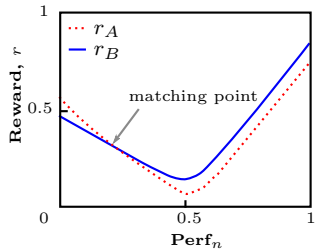


Figure 4.2.: The rising optimum reward structure in the haptic context.

there is not much support from the agent, therefore small task performance metrics are expected and the system will be easy to overrule [Abbink et al., 2012].

This trade-off way be defined as a TAFCT in which the assistance is confronted with two choices: 1) assist the user to improve nominal task performance or 2) give task authority to the user. In the former the human is 'over-assisted' while in the latter the human is 'under-assisted'. Note that it is always possible to define two situations in the assistance: one is the nominal task execution, and the other is those in which unmodeled or unexpected situations occur.

To solve this trade-off problem we incorporate a decision-making model in the haptic assistance as follows. Consider that if in the n previous samples the assistance improves nominal task performance, a high assistance level is provided. Then, nominal task performance should be satisfactory if the human agrees with the proposed assistance. Instead, if the assistance gives the control to the user and lets that him/her deviate from the nominal task, a small assistance level is provided. Then, the effort should be low and the user feels that the interaction performance increases.

Next sections present elements involved in the decision-making-based haptic assistance, namely the reward structures, the calculated evidence as metrics on human-robot performances and the decision rules.

4.2. The rewad structures

In cognitive experiments different reward structures are designed to expose certain human behaviors in decision-making for TAFCTs (e.g. exploit or explore). In our work, the reward structure determines the haptic assistance behavior. These functions are calculated based on human-machine interaction metrics; metrics which are weighted by the rewards.

The reward structure may provide a wide range of assistance behaviors. For instance, let's explore *the rising optimum* reward structure in the haptic scenario to illustrate the relationship between performances and rewards (Fig. 4.2). If a performance takes a value of 0.5 ($\text{Perf}_n \approx 0.5$), the reward associated with this performance is a global minimum value ($r \approx 0.125$), which punishes this state. Considering r_B as the reward defining the comfort when executing a task and assuming the actual comfort is a half of the maximum value ($\text{Perf}_n = 0.5$), then the reward has a value approximately of 0.25 ($r(0.5) \approx 0.25$), but when the comfort is small, the reward has a value of approx. 0.48 ($r(0) \approx 0.48$) i.e. $r(0.5) < r(0)$. This means that the assistance accumulates rewards when the operator experiences no comfort and penalizes it when s/he experiences a half of the maximum.

As humans adopt the *matching point* strategy, the decisions making models considered here also reproduces this behavior. Therefore, the decision making model-based assistance avoids low rewards given near (and in) the global minimum presented in the rising optimum structure. Actually, humans (and the models) try to maximize the total reward, most of the humans settle in the matching point. This means that incorporating the rising optimum structure restricts the performance values of $r_{A,B}$ near the minimum, assigning a minimum reward.

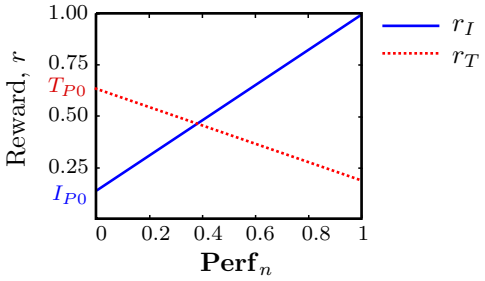


Figure 4.3.: The matching shoulder task with linear reward schedules

4.2.1. Matching shoulder reward structure

Although there are several reward functions (r_A and r_B) that have been proposed in cognitive science (Fig. 2.7), a *matching shoulder* structure (Fig. 4.3) is proposed to model the trade-off in haptic assistances, mainly because

- Eventhough the dynamics between task and interaction performance are non-linear in agreement and disagreement situations, a linear approximation reduces the complexity of the model [Passenberg et al., 2013].
- Experimental results show that if the user's plan agrees with the proposed assistance, task performance was found to be improved with increased assistance levels. For the non planning targeting task, i.e. unexpected situations, this relationship was inverted [Marayong and Okamura, 2004].
- The matching shoulder structure imposes a linear relationship between performances and reward, making it intuitive to represent different ihuman-robot types of behavior in the haptic assistance scenario (cf. Sec. 4.5.1 page 40).

The rewards in the two trade-off situations may be calculated by two performance metrics as follows: 1) a T metric is a function of how good the nominal task is performed and 2) a I metric is a function of the physical effort when unmodeled situations occur, e.g. establishing if the user is fighting against the proposed assistance.

In cognitive science experiments, the reward is calculated based on the allocation of the choice "A" i.e. the history of the choices made by the human. In decision-making-based haptic assistance, the reward is calculated based on measures of the performance, which establishes the dynamics of the interaction between the human and the assistance. A calculation of the reward with $\text{Perf}_n \in [T_P\ I_P]$ is proposed, where T_P is the task performance and I_P is the interaction performance.

Assume that only two choices are available, as in TAFCT (i.e. namely T and I), and that the reward $r(t)$ is given by

$$r(t) = \begin{cases} r_I(I_P) & \text{if } z(t) = I, \\ r_T(T_P) & \text{if } z(t) = T, \end{cases} \tag{4.1}$$

where $z(t) \in [T, I]$ is the decision made at time t. The functions r_T and r_I represent the rewards for each choice, T_P and I_P are performance measures chosen to evaluate the interaction human-robot. The reward structure is given by a matching shoulder set of equations,

$$\begin{aligned} r_I &= k_I I_P + I_{P0}, \\ r_T &= k_T T_P + T_{P0}, \end{aligned} \tag{4.2}$$

where k_T and k_I are the ratio of the change (slope) between performances and rewards I_{P0} and T_{P0} are the reward at the lowest and highest performance, respectively.

Even though we have defined the reward functions that model the trade-off between the chosen performances, the reward does not give any method on how the decision $z(t)$ maximizes the total intake. To this end, we consider the Win Stay Lose Switch rule known as γ-policy and the Drift Diffusion (DD) model, proposed in cognitive science literature.

4.3. Performance metrics

Several qualitative and quantitative performance measures can be considered when designing haptic assistances. Amongst common performance measures that could be listed we have [Passenberg et al., 2013]: task performance, agreement, smoothness, workload-effort and/or efficiency.

The main objective is improving task performance without degrading human experience, but these metrics depend on the objectives of the particular task. For example, a successful assembly task requires to determine whether the object was assembled correctly, the forces indicating physical effort or possible damage of the objects; when a tracking task is considered, the position error estimates the success of the task; or in human-robot cooperative transport tasks, internal forces indicate collaborative strategies.

Other metrics used in the haptic literature are [Passenberg et al., 2013]: the smoothness measure which relates to the jerky movements induced by the assistance and may affect task performance; the effort or workload which depends on the human physiological and mental state (the effort is calculated by the forces, energy or physiological signals give indications of the user's physical or cognitive workload); and efficiency which relates performance and effort (a small effort and a high performance results in high efficiency).

In cases when unmodelled situations occur or when a multiple performance criterion is necessary to maintain good task execution, a mechanism that solves the trade-off between different performances is required. Furthermore, if the assistance-mechanism only includes one task performance as the objective function, humans can become dependent on assistive guidance and s/he may be poorly prepared for unexpected situations [Powell and O'Malley, 2012].

The selected measures for task performance and interaction performance in this thesis can be considered generic examples, but can be customized when applying the approach to a specific application.

As a tracking task with un-modeled obstacles is considered, we calculate the reward based on the tracking error as nominal task performance measure (T_P), and the internal forces as an agreement performance measure (I_P).

4.3.1. Tracking path performance

The performance measures are normalized in the perpendicular direction of the path over an observation window of n samples. In order to normalize the measurements, the maximum value is updated and calculated for each window.

The tracking path performance is considered as the **nominal task performance**, which is measured via the absolute position error between the desired path and the current position of the master; the normalized version of the tracking performance measure is given by

$$T_{P,n} = \frac{mean_n(\|\mathbf{e}\|)}{\max(\|\mathbf{e}\|)},$$

where $\mathbf{e}$ is the position error.

Note that low task performance is rewarded with high payoff; while high task performance is rewarded with a low payoff (cf. Fig. 4.3).

4.3.2. Agreement performance

Internal forces occur if two partners push or pull in different directions and may be used as an interaction measure [Passenberg et al., 2013]. 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 [Groten, 2011].

The (haptic) interaction performance is defined as the agreement measurement between the operator and the master device. Agreement is related to the internal forces between human and haptic device. Internal forces are defined by [Groten, 2011]

$$\mathbf{f_i} = \begin{cases} \mathbf{f}_\perp & \text{if } sign(\mathbf{f}_\perp) \neq sign(\mathbf{f_a}_\perp) \wedge \|\mathbf{f}_\perp\| \leq \|\mathbf{f_a}_\perp\| \\ -\mathbf{f_a}_\perp & \text{if } sign(\mathbf{f}_\perp) \neq sign(\mathbf{f_a}_\perp) \wedge \|\mathbf{f}_\perp\| > \|\mathbf{f_a}_\perp\| \\ \mathbf{0} & \text{if otherwise} \end{cases}, \tag{4.3}$$

where $\mathbf{f}_\perp$ is the force exerted by the human and $\mathbf{f_a}$ is the force exerted by the assistance. All forces are measured in the perpendicular direction of the desired path. The interaction performance measure is finally defined by

$$f_{i,n} = \frac{mean_n(\|\mathbf{f_i}\|)}{\max(\|\mathbf{f_i}\|)}, \qquad I_{P,n} = 1 - f_{i,n}, \tag{4.4}$$

where $I_{P,n}$ is the interaction performance and is referred to as an agreement measure; $f_{i,n}$ is the normalized version of the internal forces where high internal forces mean low degree of agreement. Low agreement is rewarded with low payoff, while high agreement is rewarded with a high payoff (cf. Fig. 4.3).

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 ($1 - f_{i,n}$). In this case, the reward r_I is defined in (4.2) as $k_I = 0.5$ and $I_{P0} = 1$.

So far we have introduced all the elements present in the decision making process and have established a model for the haptic assistance. The reward structure represents the relation between performances and rewards; the reward in turn provides evidence to decide about the assistance levels. We also introduced performance measures which may be used to implement the assistance and that represent the trade-off present in the task proposed, i.e. tracking error (nominal task) and agreement (unmodeled situations).

We are now ready to present the decision making model and its dynamics when the human interacts with the assistance in both modeled and unmodeled situations. This is interpreted as movements following the nominal task and apart from it, respectively.

As the haptic assistance is implemented as a guiding virtual fixture (GVF), the parameters are derived via experimentation in both circumstances, a high assistance level to reduce the nominal task error and a low assistance level to reduce the disagreement when the user avoids the obstacles.

The assistance was implemented in the teleoperation system presented in Sec. 3.4.

4.4. γ-policy in haptic shared control

The Win Stay Lose-Switch (WSLS) is the simplest decision making rule and it has been is extensively used in cognitive science. The γ-policy is a WSLS-like decision making rule.

The γ-policy-based assistance (Fig. 4.4) supports the human to improve nominal task performance or interaction performance when unmodeled situations occur. The $\gamma-$policy is defined as,

$$z(t+1) = \begin{cases} z(t) & \text{if } r(t) > r(t-1) \\ \sim z(t) & \text{otherwise} \end{cases} \tag{4.5}$$

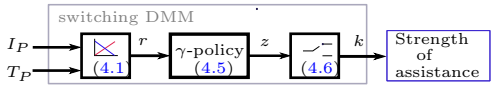


Figure 4.4.: The γ-policy based assistance consists in a switching policy between two assistance levels; the policy depend on the last reward (r). Each block shows its corresponding equation number, namely: The reward structure (4.1), the γ-policy (4.5), and the switching allocation of the assistance level (4.6)

where $z(t)$ is the decision made at time t ($z(t) \in [I\ T]$), $r(t)$ is the reward at time t and the symbol $\sim$ denotes the "not" operator; i.e. if $z(t) = T$ (resp. $z(t) = I$), then $\sim z(t) = I$ (resp. $\sim z(t) = T$).

The 'greater than' sign in the $\gamma-$policy (4.5) means that the assistance keeps assisting the human in the current choice unless the reward is decreased. The original WSLS rule admits 'greater than or equal' sign.

Note that only the last reward is taken into account to decide about the suitable assistance level, thus no history (or memory) on previous rewards is considered.

Next, a γ-based gain scheduling strategy is proposed.

4.4.1. γ-policy-based gain scheduling

The human is assisted in the task execution by adapting the assistance level in a binary fashion; the assistance-mechanism switches between two levels as follows,

$$k = \begin{cases} k_{high} & \text{if } z(t) = T \\ k_{low} & \text{if } z(t) = I \end{cases} \tag{4.6}$$

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

This haptic assistance tries to emulate the switching assistances strategies reported in [Kragic et al., 2005, Oguz et al., 2010, Passenberg et al., 2013]. The dynamics of the $\gamma-$policy in a haptic scenario is presented as follows,

$\gamma-$policy response

To explain the assistance strategy, the interaction is divided into two types of movements: on and off the path. When the human moves on the path the samples refer to nominal task execution, while when the human moves off the path the samples refer to unexpected situations (the obstacles).

In movements on the path, the position error ($T_{P,n}$) is low and overall the agreement ($I_{P,n}$) remains high. Therefore, the reward ($r(t)$) is high. In this case, the assistance should select a high stiffness to assist the human movements on the path, i.e. choosing option T ($k_\perp$ in Fig. 4.5).

The γ-policy presents a chattering behavior in the decision making process ($z(t)$ in Fig. 4.5). This chattering occurs because the system switches again when the reward decreases (even in small amounts). This situation occurs both on-path and off-path movements.

In general, this behavior occurs because the γ-policy makes decisions taking into account only the last reward. This behavior is depicted for reward ($r(t)$) in the movements off the path in Fig. 4.5.

4.5. Drift Diffusion model in haptic shared control

In the Drift-Diffusion (DD) model-based assistance architecture (Fig. 4.6), the decision-making model varies the strength of the assistance based on the probability to improve certain human-system performance variables (T_P or I_P) depending on the choice made. This probability is

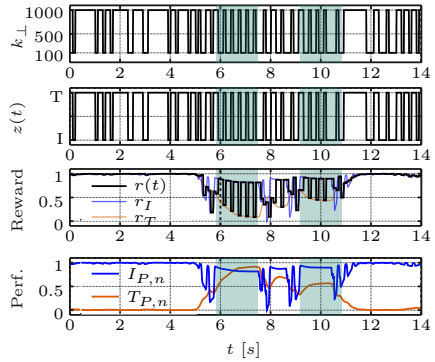


Figure 4.5.: γ-policy based assistance behavior. $k_\perp$ is the assistance level which corresponds to the stiffness in the perpendicular direction of the desired path; r_I and r_T are the rewards for interaction and task performances, respectively; $r(t)$ is the reward; $T_{P,n}$ and $I_{P,n}$ are the normalized version of task and interaction performances, respectively; and z is the decision making, I for interaction or T for task performance.

calculated as a function of the accumulated evidence (Δ) that serves as a mechanism to decide whether to follow the plan for the nominal task (T_P) or to give the control to the user (I_P). The decision-making mechanism accumulates evidence based on the actual and past rewards ($r(t)$). When the accumulated evidence decreases below a certain threshold, the DD switches to the other choice (z), calculates the evidence of the new choice and keeps this decision until the accumulated evidence decreases again below a certain threshold.

The DD model-based assistance includes information on the reward obtained in each decision in order to accumulate evidence (w) for the decision making process. In order to take this information into account, an update rule for the evidence in favor of each choice is proposed by,

$$\begin{aligned} w_z(t+1) &= w_z(t) + \lambda(r_z(t) - w_z(t)), \\ w_{\sim z}(t+1) &= \lambda r_{\sim z}(t), \end{aligned} \tag{4.7}$$

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.

The DD model is a reinforcement learning model in which evidence in favor one choice over the other is integrated until a predetermined threshold is reached; the *soft-max* model is used as the DD model strategy such that the probability of choosing the improvement of the interaction performance (I) over the task performance (T) (or vice versa if desired) is given by

$$p_I(t+1) = \frac{1}{1 + \exp^{-\mu(w_I - w_T)}}, \tag{4.8}$$

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.

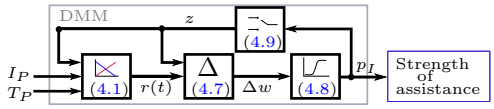


Figure 4.6.: Diagram block of the DD model-based assistance. Each element shows its corresponding equation number, namely: The reward structure (4.1), a reinforcement rule to accumulate the evidence (4.7), the drift-diffusion model (4.8) and the decision-making rule (4.9).

Larger μ represents a high slope in the sigmoidal function (4.8), *this means that exists more certainty in the decision making process* Fig. 2.8.

Equation (4.7) is inspired by the reinforcement rule presented in [Bogacz et al., 2007] 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 was not chosen ($\sim z$) is updated without any learning rule, just with the previous reward obtained. The parameter λ in the accumulation of the $\sim z$ choice, scales the reward to maintain no preference over any option when interaction commences.

The accumulated evidence can be seen as the memory of the model on previous rewards. More memory means that the model takes into account more history to make the decision. The $\lambda-$parameter in (4.7) represents the amount of memory in the DD model. If λ is low, the model takes into account more history, this is, more memory. On the other hand, larger λ means less memory.

The evidence w_z is accumulated based on the decision z which is given by

$$z(t+1) = \begin{cases} I & \text{if } p_I > 0.5 \\ T & \text{otherwise.} \end{cases} \tag{4.9}$$

4.5.1. DD model-based gain scheduling

The haptic assistance level (high or low gain) depends on the probability (4.8) that a certain choice improves the total system reward.

$$k(p_I) = \frac{1}{2} + \frac{1}{2}\tanh\left(\frac{p_I - \phi}{\varphi}\right) \tag{4.10}$$

where ϕ and φ are user defined parameters; ϕ is the switch point and φ is the smoothing level.

An alternative consist of mapping the probability ($\alpha(p_I)$) to the level of the haptic assistance via a linear homotopy function

$$k = \alpha\ k_{low} + (1-\alpha)\ k_{high} \tag{4.11}$$

where k is the assistance level.

Observe that even though the assistance decides on one of the two alternatives, the assistance level is built on the probability of the improvement of the interaction performance.

DD model response

The dynamics of the DD model-based assistance depend on the accumulated evidence ($\Delta w = w_I - w_T$) in favor of assisting to improve a certain performance. When the agreement between the human and the assistance decreases, the assistance level is reduced. Therefore, the operator is allowed to retain full authority over the task. On the other hand, when the agreement increases, the assistance level is increased. Therefore, the nominal task performance increases.

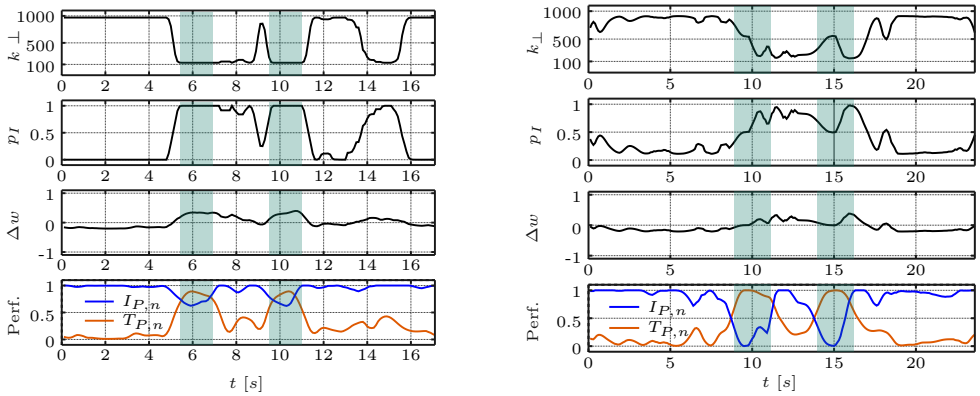
Different model dynamics are presented by choosing different values of the parameters μ and matching shoulder structure. The λ parameter is selected to be constant in all the experiments in order to reduce the number of parameters to be analyzed simultaneously.

Marked periods in Figs. 4.7 and 4.8 indicate phases in which the operator moved off the path.

When the operator moves off the path, the position error ($T_{P,n}$) increases and the agreement ($I_{P,n}$) decreases (Perf. in Fig. 4.7a and 4.7b).

If μ is large, the operator experiences sudden changes. Because of this, the user experience is degraded. Hence, the parameter $\mu = 10$ enables a smoother transition between assistance levels (Fig. 4.7b with $\mu = 10$).

The assistance strategy must be defined considering the "specific" objectives of the task; for example, telesurgery might require to restrict the movement of the operator to a given path. In



(a) $\mu = 100$ A larger μ-parameter the user experiences a sudden change on stiffness

(b) $\mu = 10$, A smaller μ-parameter a smoother transition on stiffness

Figure 4.7.: Drift-Diffusion-Model behavior for two μ-parameter values and $\lambda = 0.7$ in a haptic tracking path task, cf. (4.8). The matching shoulders dynamics without any preference on performances is presented

other tasks path tracking may be less important, for example, when the operator is required to avoid 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} (4.2) adjust the preference to assist the operator in the task or the interaction.

The preference to assist the operator in tracking may be achieved by considering $k_T = -1, k_I = 0.5, T_{P0} = 1, I_{P0} = 0$. The maximum achieved reward for interaction is half of the one set in the previous configuration, giving more importance to task performance rather than to interaction performance. Note also that lower interaction ratio (k_I) compared with task ratio (k_T) means that the accumulated evidence is reduced for interaction if it is compared to task, which explain the preference for task performance. Likewise to weight the interaction, the reward for the task performance can be reduced; for example, considering $k_T = -0.5, k_I = 1, T_{P0} = 0.5$ and $I_{P0} = 0$.

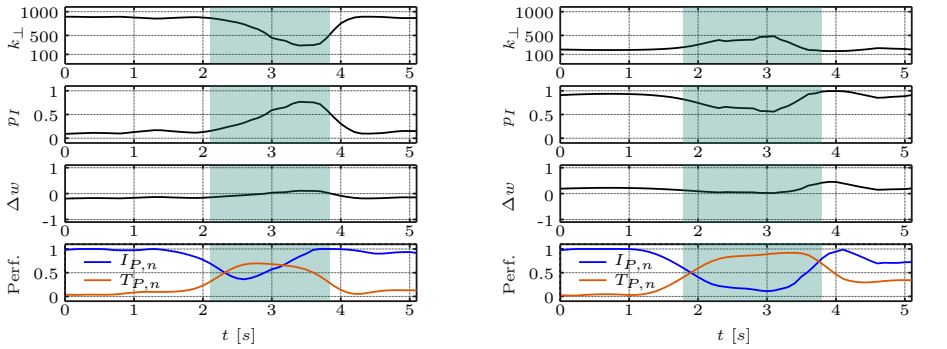
When task performance is weighted in favor of interaction performance, the evidence (Δw) to choose a stiffness that improves interaction performance (p_I) is low, therefore the probability is low (Fig. 4.8a). In contrast, when the interaction is weighted in favor, the evidence to assist in the interaction is high (Δw), so the probability to assist the operator to improve the interaction is also high (Fig. 4.8b).

4.6. Conclusion

The chattering in γ-policy rule-based assistance produces oscillations in the assistance level (the stiffness), which degrades the user experience. The main drawback in the γ-policy for haptic assistance is that it just considers the last reward. Thus when a little change occurs in the performances, the model cannot incorporate memory to keep choosing a strategy that enhances a certain performance. This means the policy does not capture accurately the decision-making dynamics that an assistance-mechanism requires in the human-machine haptic scenario.

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.

In next chapter, 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-like and no-assistance strategies.



(a) Task pref., $k_T = -1, k_I = 1, T_{P0} = 1$ and $I_{P0} = 0$, The evidence (Δw) to choose a stiffness that improves interaction performance (p_I) is low

(b) Interaction pref., $k_T = -0.5, k_I = 1, T_{P0} = 0.5$ and $I_{P0} = 0$
The evidence to assist in the interaction is high (Δw), so the probability to assist the operator to improve the interaction is also high

Figure 4.8.: Response for the Drift-Diffusion model-based assistance with two different matching shoulder conditions, cf. (4.2).

5. Methods

5.1. Main Hypothesis

The objectives of the experimental analysis are two-fold: the first is concerned with the decision-making model itself and seeks to validate if the proposed assistance is successful; the second seeks to establish whether different reward structures describe accurately different problem situations, and if they actually help the operator to execute the task at hand. In order to achieve this, we collected objective data that presents the subject and the DD model-based assistance behavior. Despite the subjective perception is not our main concern, we ask the users two questions to present insights on the user perception about the assistance conditions (cf. Sec. 5.7.2)

To evaluate the performance of subjects and DD model-based assistance, the main hypotheses that we tested were,

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

H2. DD model-based assistance performs similar to human in terms of task performance, interaction performance, physical dominance and physical effort.

H3. DD model-based assistance improves efficiency compared with Na and Ha.

H4. The DD model-based assistance reproduces different assistance behaviors by means of the reward structure setup, as follows,

Tracking error: $Tk < nP < In$
Internal forces performance: $In < nP < Tk$
Physical dominance: $In < nP < Tk$
Physical effort: $In < nP < Tk$
Efficiency: $Tk \leq Ip < nP$

where Tk, nP and In represent the reward setup for task preference, no preference and interaction preference, respectively.

This chapter presents the methods used in a user study performed to evaluate the proposed assistance; which is organized as follows: first, a summary of the conducted statistical test is presented. Next, the experimental setup, conditions, apparatus, participants, procedure and measures are shown; finally a questionnaire are introduced which gives useful information on how the assistance strategy is perceived by the user.

5.2. Statistical tests

The user study refers to statistical comparison of the different assistance strategies. We conducted an analysis of variance (ANOVA) which is a common statistical test performed in haptic assistance studies.

As the particular test depends on assumption on the normality of the data, the normality on objective measures is evaluated by the Kolmogorov Smirnov (KS) test (commonly also use in haptic studies). We conducted a *Friedman test*, which is a non-parametric analysis of variance (ANOVA)—because we cannot assume normality w.r.t. KS test—except for efficiency measure. All statistical tests were performed with a 5 % level of significance.

The analysis also includes an effect size test (r) which indicates the magnitude of the effect caused by a factor (a factor in our experiment refers to different condition of assistance). The effect size is independent of the sample size thus, the effect size can complement some of the shortcomings of the Null Hypothesis Significance Testing and p-value analyses presented in the ANOVA test. The effect size is calculated for the post-hoc comparisons. The effect size may consider the normality of the samples (e.g in standardised mean effect size); For this purpose the exact Wilcoxon Mann-Whitney rank sum test is used. Although the interpretation of the size depends on the context of study, a general guideline is shown in Table 5.1.

Table 5.1.: Standard value for the effect size (r) [Cohen, 1988]; it is used also in Pearson's coefficient of correlation.

	small size	medium size	large size
|r|	0.1	0.3	0.5

Further, we investigate the correlation between physical dominance and efficiency. Therefore, we performed a Pearson's rank correlation test. The null hypothesis of this test states that no correlation between the two variables exist. The correlation value represents the strength of the linear-relationship between variables, which is represented as a value (r) between -1 and 1. The variables have to be symmetrical i.e. there is no notion that one of them precedes or causes the other. Otherwise, if the variable is not symmetrical, a regression analysis would show how one variable explains the other (one variable is controlled and the other is explained in terms of the first). We noted the correlation by r because the Pearson's test represents an effect size test.

5.3. Experimental Setup

The operator was asked to follow a desired path which was drawn on a piece of paper placed on the remote site and within the workspace of the slave (Fig. 5.1). 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. 5.1b). No specific preference was induced for either of the two options.

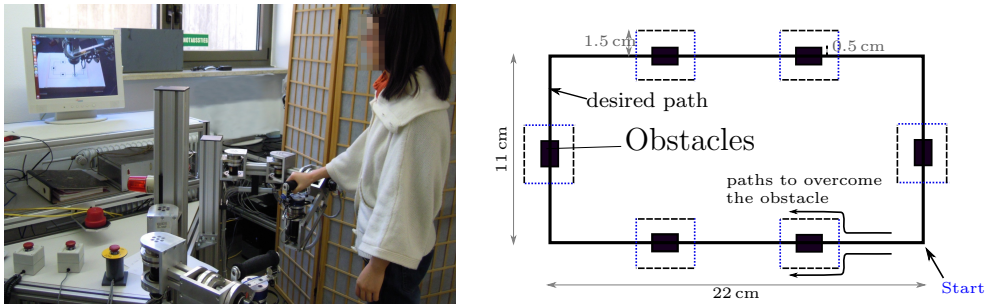
Visual feedback was provided to the user by a video link that transferred a mono camera image from the remote site and displayed it via a LCD screen placed at master side.

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.

5.4. Conditions

As the principal goal is the evaluation of the DD model in the haptic assistance scenario, three different conditions were considered that explore the influence of the matching shoulder reward structure parameters in the haptic assistance, cf. (4.2); to compare the DD model-based assistance two additional conditions are proposed, a human-like assistance and no-assistance condition:

1. $\boldsymbol{nP}$, the decision-making model has no preference for any performance measure, $k_T = 1, k_I = 1, T_{P0} = 1, I_{P0} = 0$;
2. $\boldsymbol{Tk}$, task performance is favored, $k_T = -1, k_I = 0.5, T_{P0} = 1, I_{P0} = 0$



(a) A user performing an evaluation trial in a pilot test. The image shows the master and the video link showing the slave environment, in which the desired path is drawn.

(b) Detail of the desired path drawn in the slave environment. The solid line indicates the nominal task; two suggested path to overcome an obstacle are shown.

Figure 5.1.: Experimental setup for the haptic task.

3. ***In***, interaction performance is favored, $k_T = -0.5, k_I = 1, T_{P0} = 0.5, I_{P0} = 0$.
4. ***Ha***, a human-expert operator haptically coupled with the tested subject.
5. ***Na***, one operator performing the task alone and without any assistance.

Although human-human interaction in a teleoperation task (Ha) may produce larger position errors compared to the experienced in natural human-human interaction, in this research we proposed such a comparison as it presents a reference for designing more natural interaction strategies [Groten, 2011].

The concept of a "natural" haptic interaction has not been clearly defined yet. A simple and accepted concept is that natural interaction occurs when the human operator performs the task directly [Powell and O'Malley, 2011]; when an assistive force is added, the operator has to adapt to this new external signal which may feel *unnatural*. However, incorporating such assistive forces improve overall performance of task execution [Kucukyilmaz et al., 2013, Mörtl et al., 2012, Unterhinninghofen, 2009, Abbink et al., 2012, Passenberg et al., 2013], hence it is of interest to find a suitable adapting mechanism between task performance and the comfort that the operator perceives. In the interaction analysis conducted in this thesis, we propose to study the interaction between the assistance and the operator as well as the interaction between humans, with the idea of compare objective interaction measures on how two humans adapt to perform the task and how natural the assistance is presented.

Condition Ha is particularly useful for studying the vary assistance mechanism because of a human expert varies the assistance level in a intuitive fashion. Besides, this experimental setup is motivated due to the benefits of implementing haptic assistances compared with no-assistance have been already established in haptic literature (cf. Chap. 2, e.g. [Marayong et al., 2002, Gunn et al., 2009, Unterhinninghofen, 2009, Abbink et al., 2012, Nitsch et al., 2012, Passenberg et al., 2013, Boessenkool et al., 2013]). Besides a guidelines for implementing varying assistance—and the comparison to constant assistance—have been already presented in [Passenberg et al., 2013]. Nevertheless, in literature a nominal-artificial varying assistance scheme is missing which enable us to compare unequivocally the proposed varying assistance mechanism.

5.5. Apparatus

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 [Kron et al., 1999].

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. 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-sensors (JR3) was mounted at the tip of the manipulators to measure interaction forces with the human operator or the environment.

To determine the suitable virtual fixture admittance a task-dependent process is required [Kragic et al., 2005]. The upper and lower bounds are calculated based on experimentation on both circumstances: a high level of assistance to reduce the nominal task error ($k_{high} = 1000\,^{\mathrm{N}}/_{\mathrm{m}}$) and a low level of assistance to reduce the disagreement when the user avoids the obstacles ($k_{low} = 100\,^{\mathrm{N}}/_{\mathrm{m}}$). These parameters are presented as bounded values because the DD-model adjusts the actual assistance strength based on p_I (4.8). The assistance is implemented using the following parameters,

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

We take advantage of the bimanual telerobotic interface to implement a human-like assistance strategy (Ha). A Multiple Operator Single Robot (MOSR) scheme is implemented, Fig 5.2. The MOSR enables two humans, both operating a human-system interface to control collaboratively one telerobot (slave) in a shared-control mode [Feth et al., 2009b].

Generally the MOSR-scheme is implemented for student-teacher scenarios, where a trainee is supported by a trainer in performing manipulation tasks [Feth et al., 2009b]. In our approach the trainer play the varying assistance role.

The admittance parameters (Y) for the MOSR scheme are equal and constant in all directions, Fig 5.2, as follows,

$$m = 5\,\mathrm{Kg}, \quad b = 10\,^{\mathrm{Ns}}/_{\mathrm{m}} \tag{5.1}$$

Note that in Na condition and DD model-based assistance conditions the human-expert is absent.

5.6. Participants and procedure

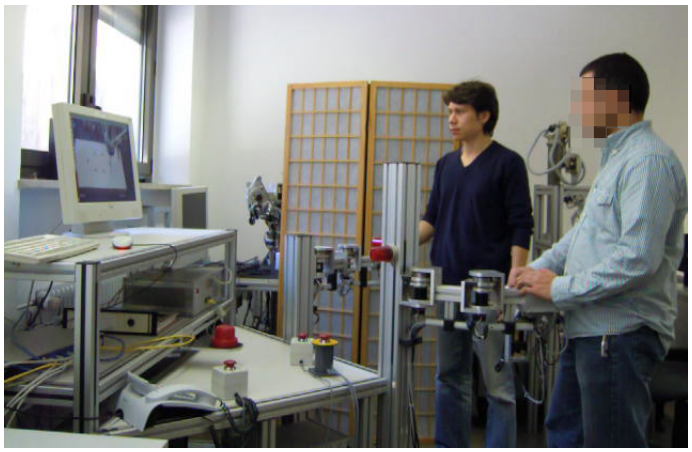
5.6.1. Participants

Twelve healthy subjects were tested and asked to perform the same task under the five aforementioned conditions.

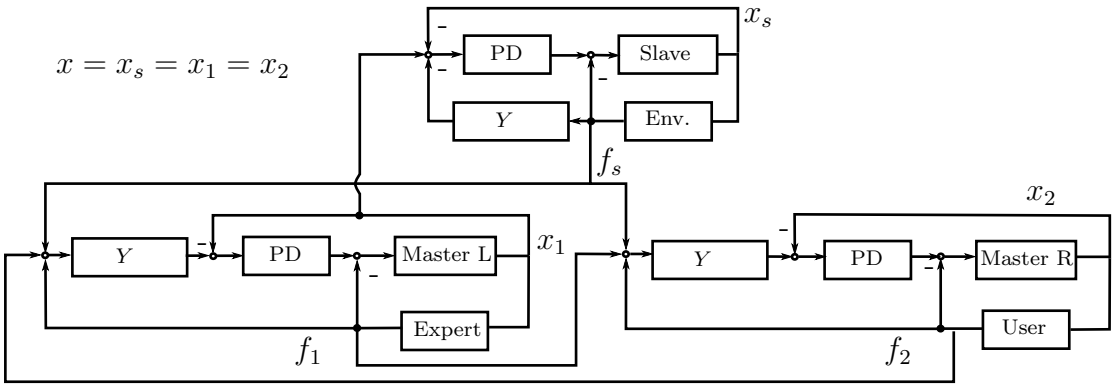
The expertise level of the users may change the outcome of the experiment due to their ability to adapt to different situations and levels of assistance. Hence, in order to reduce this bias unexperienced test subjects are also studied. An expert user though takes the role of the haptic assistance in the human-human condition. For instance, an inexpert user might assume a conservative role during task execution, while more experienced users can take a more active role on task execution. The expert user in Ha assists each of the users based on the requirement of each subject, this may be seen as the natural varying assisting policy.

5.6.2. Procedure

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



(a) A dyad performing an evaluation trial in a pilot test. The image shows the bimanual teleoperator, which consist of the left master (Master L) and right master (Master R). A video link shows the slave environment, in which the desired path is drawn.



(b) The MOSR control architecture

Figure 5.2.: Multiple Operator Single Robot (MOSR) scheme, two humans jointly perform a tracking task. A partner adopts the assistance role, this partner is called the expert. The MOSR scheme implements the human-like assistance (Ha) condition in our experiment

passive role in the execution of the task. This consideration is relevant because in teleoperation the user has to maintain aware of task execution. At the end of each trial, the operator is asked to answer two questions (Sec. 5.7.2).

5.7. Measures

The response of DD model-based assistance is evaluated by objective metrics; the human perception is inquired via subjective measures.

We mainly adopt the experimental evaluation proposed by Groten [Groten, 2011] due to this metrics explicity consider a dyadic measures in haptic studies.

The objective measures considered are: the root-mean-square deviation (RMS) of the tracking error as task performance measure; the interaction is evaluated by both, the agreement between human and assistance, and the physical dominance on task execution; a power-based effort measure is considered to evaluate the work done by the dyad on task. We relate the effort and task performance by an efficiency measure that was proposed in a human-human haptic interaction. In addition, we correlates physical dominance with task performance and efficiency to improve the discussion on the shared strategies adopted on the experiment.

Additionally two questions were posed at the end of each experimental trial to assess the subjective feeling in terms of perceived interaction performance with the assistant and overall achieved task performance.

5.7.1. Objective measures

Tracking error

A root-mean-square of tracking error ($RMS(e)$) is considered as *task performance measure* for the interaction sequel. We analyzed movements on (nominal task) and off (unmodeled situations) the desired path separately. Thereby a better understanding for the DD model-based assistance behavior in these two situation is reported.

The RMS is defined by,

$$RMS(e) = \sqrt{\frac{\sum_{i=1}^{N} \left(x_{d,\kappa} - x_{m,\kappa}\right)^2}{N}} \tag{5.2}$$

where i is the time step ($\kappa = [1, \ldots, N]$), $x_{d,\kappa}$ is the desired position defined in the nominal task description, $x_{m,\kappa}$ the actual position of the master device and N the number of samples in the examined interaction sequence.

Internal forces

Internal forces were calculated as already defined in (4.3), page 37. The mean of internal forces describes the disagreement on each interaction sequence, mean which is denoted by f_i. Again, we analyzed movements on and off the path separately.

Physical Dominance

While high-level dominance determines which operator determines the strategy to move the object (planner action), the physical or low level of dominance determines which operator applies higher manipulation force to execute the strategy that is determined at the high level of dominance (cognitive level). Here we are interested in the physical dominance measure because we investigate

the distribution of the work forces exerted by the user and the assistance (no preference on the off-path trajectory is considered).

In teleoperation, the user have to control task execution. Thus we expected that the user determines the strategy (cognitive level) to move forward the object. Note that the user has no-assistance in the parallel direction of the path, thus the user have to control (autonomously) the task execution.

The partner in a dyad who applies higher manipulation forces commands the object movement to a higher degree and can thus be considered dominant [Groten, 2011].

Internal forces ($\mathbf{f_i}$) are important to describe the individual efforts (wasted effort) of each member of the dyad, but they cannot establish the strategies adopted by the dyad because the internal forces do not contribute to the motion of the object. On the other hand, the physical dominance of partner 1 over 2 (PD_{12}) is defined by the external force which generated work as follows [Groten, 2011],

$$PD_{12,\kappa} = \frac{\|\mathbf{f}^1_{ext,\kappa}\|}{\|\mathbf{f}_{sum,\kappa}\|}$$

where, $\mathbf{f}_{sum,\kappa}$ is the sum of external forces and $\mathbf{f}^1_{ext,\kappa}$ is the external force exerted by partner 1, which can be calculated by

$$\mathbf{f}^1_{ext,\kappa} = \mathbf{f}^1_{,\kappa} - \mathbf{f}^1_{\mathbf{i},\kappa} \tag{5.3}$$

$$\begin{aligned}\mathbf{f}_{sum,\kappa} &= \mathbf{f}^1_{ext,\kappa} + \mathbf{f}^2_{ext,\kappa} \\ &= \mathbf{f}^2_{,\kappa} + \mathbf{f}^2_{,\kappa}\end{aligned} \tag{5.4}$$

where, $\mathbf{f}^1_{,\kappa}$ is the total force exerted by partner 1 and $\mathbf{f}^1_{\mathbf{i},\kappa}$ is the internal force exerted by partner 1, and (5.4) implies that $\mathbf{f}^1_{\mathbf{i},\kappa} \equiv -\mathbf{f}^2_{\mathbf{i},\kappa}$. Note that the physical dominance is calculated in each sample (κ); thus in an interaction sequel, the dominance may be calculated by its mean.

A partner is absolutely dominant with a value of one, and absolutely non-dominant with a value of zero. $PD_{12} \in [0, 1]$ 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 [Groten, 2011],

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

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 value unequal zero 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 object. Both measures complement the analysis in a jointly dyadic haptic interaction.

Power-based effort

Higher energy flow relates to a higher physical effort. Effort measure is calculated as a function of the power exchanged on the interaction sequence,

$$P_* = \dot{\mathbf{x}}\ \mathbf{f}_* \tag{5.5}$$

where, P_* is the power from partner '$*$' to the environment, $* \in [1, 2]$, $\dot{\mathbf{x}}$ is 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.

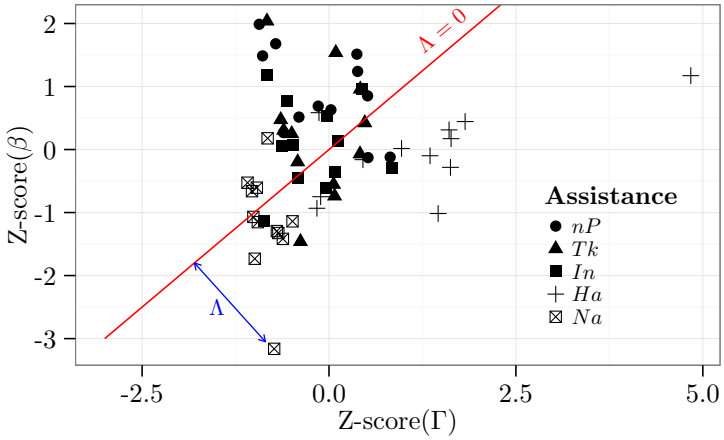


Figure 5.3.: Dyadic efficiency (Λ) as function of performance (Z-score(β)) and effort (Z-score(Γ)) for the subjects in the haptic experiment. Points below the diagonal represents inefficiency (more effort and less performance) behavior. While, points above or in the diagonal represents an efficient dyad (less (or equal) effort for more (or equal) performance)

A positive energy flow, i.e. energy injection to the system (e.g. acceleration of the virtual object), causes physical effort for the operator, also a negative energy flow conduces to physical effort, i.e. dissipating energy from it (e.g. deceleration of the virtual object).

The effort measure on a dyadic interaction represents the total effort used in each assistance type and can be calculated by [Groten, 2011],

$$\Gamma = MAP^1 + MAP^2 = \frac{1}{N}\sum_{\kappa=1}^{N}|P^1_{,\kappa}| + \frac{1}{N}\sum_{\kappa=1}^{N}|P^2_{,\kappa}| \tag{5.6}$$

where, $P^1_{,\kappa}$ and $P^2_{,\kappa}$ is the energy flow at the respective interfaces points at a given time step κ. Superscripts $1, 2$ over MAP indicate measurement for the partner 1 or 2 and the total-dyadic level is represented by Γ.

Based on the force ($\mathbf{f}$) used to calculate the power, we may study different aspects on the interaction. For instance, if we use the internal force in (5.5), we would study the physical effort in disagreement situations; if we used instead the total force, we would study the total effort in the interaction sequel. We analyzed the total force and its component exerted on the perpendicular direction of the desired path, for both cases the individual and dyadic power.

Efficiency

An efficiency measure expresses the relation between performance and effort. Efficiency is high when high performance is gained with low effort. Efficiency (Λ) may be described in two-dimensional space with a performance-axis (y-axis) and an effort-axis (x-axis), where the two measures are z-score standardized[1] (mean = 0, std. deviation, = 1) to accommodate differences in measurement scales [Groten, 2011].

A reference line where $\Lambda = 0$ is defined by the linear function, Performance = Effort (both z-scored), cf. Fig. 5.3. Any particular observation of effort and performance defines a point in this

[1] Z-score(x)= $\frac{x-\mu}{\sigma}$, where μ is mean, σ is the std. deviation and x is the data

space, and the corresponding efficiency can then be calculated by the perpendicular distance of the point to the reference line ($\Lambda = 0$).

On haptic collaboration the efficiency measure may be expressed by [Groten, 2011],

$$\Lambda(\beta, \Gamma) = \frac{\text{Z-score}(\beta) - \text{Z-score}(\Gamma)}{\sqrt{2}}$$

where Z-score(β) is a z-standardized performance measure and Z-score(Γ) a Z-standardized effort measure.

The RMS of tracking error is chosen as performance measure (5.2). Because in this efficiency definition, performance values are positively defined, the RMS error is transformed to receive a positive measure (i.e. high values mean good performance) as follows [Groten, 2011]:

$$\beta = 1 - \frac{RMS_j}{RMS_{\max}}$$

where $RMS_{\max}$ is the maximum mean of error found in a given data set (maximum of the whole experimental data, $RMS_{\max} = 0.0044\,\text{m}$) and RMS_j is the RMS error of trial j.

Effort (Γ) is expressed as power-based measure in the dyadic case (5.6). Positive/negative efficiency values describe efficient/inefficient behavior [Groten, 2011].

5.7.2. Subjective perception

The following two questions were posed at the end of each experimental trial to assess the subjective feeling in terms of perceived interaction performance with the assistant and overall achieved task performance:

1. In which trial did you feel that the task was best executed? (Task performance)
2. In which trial did you feel that the task was easier to execute? (Interaction performance)

The answers were mapped to one of the five conditions presented: nP, Tk, In, Ha, and Na.

6. Results and Discussion

6.1. Objective measures results

6.1.1. Tracking position error

We cannot assume the normality for RMS of tracking-position error based on KS test[1]. Therefore, a no-parametric ANOVA test is performed for the RMS value of the tracking-position error in which samples for on and off path movements are analyzed separately.

On-path tracking error position

A Friedman test revealed a significant effect of assistance strategies on tracking error when the user moves on the path ($\chi^2(4) = 32.267$, $p < 0.001$[2]). All post-hoc test were conducted using a Wilcoxon test with Holm-Bonferroni correction. Significant differences between Na and all other assistance strategies were found $\{nP, Tk, Ha\}$ ($p < 0.05$, $r = \{0.778, 0.849, 0.66\}$), except In condition ($p > 0.05$). Differences are established between nP and $\{In,\ Ha\}$ ($p < 0.05, r = \{0.436, 0.495\}$). Besides, differences were found between Tk and $\{In, Ha\}$ ($p < 0.05, r = \{0.731, 0.801\}$). In contrast, no signficant differences were found between In and $\{Ha, Na\}$ ($p > 0.05, r = \{0.13, 0.53\}$), nor between nP and Tk ($p > 0.05, r = 0.33$). The differences are highlighted in Table 6.1.

Table 6.1.: p-value (p) and effect size (r) for RMS value of the tracking-position error when subjects move on the path.

	nP		Tk		In		Ha	
	p	r	p	r	p	r	p	r
Tk	0.185	0.330	-	-	-	-	-	-
In	**0.007**	0.436	**0.005**	**0.731**	-	-	-	-
Ha	**0.049**	0.495	**0.005**	**0.801**	0.622	0.130	-	-
Na	**0.015**	**0.778**	**0.005**	**0.849**	0.081	**0.530**	**0.017**	**0.660**

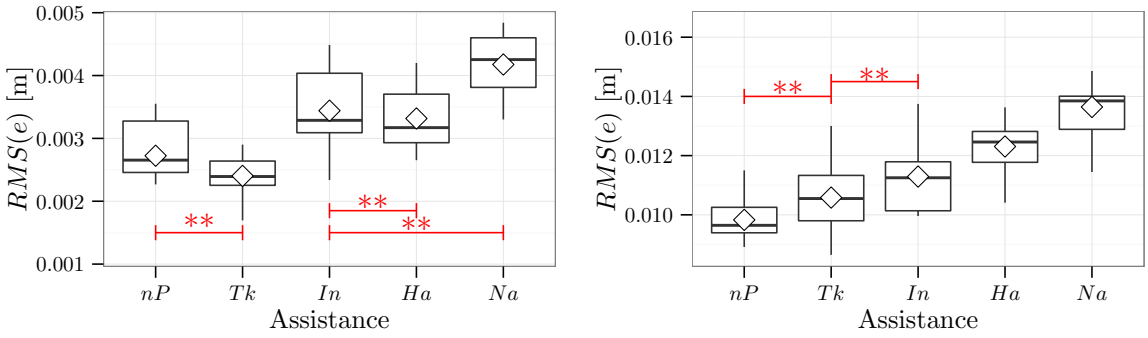
Regarding the experimental evidence on task performance, when the user moves on the path (Fig. 6.1a and Table 6.1), the DD model-based assistance conditions improve task performance compared to Na ($p < 0.05$), except when In is compared with Na. In this case no statistical difference was found ($p = 0.081$).

When the user moves on the path, the DD model-based assistance improves nominal task performance compared to Ha ($p < 0.05$); but no significant difference was found between In and Ha ($p = 0.62$).

Comparing the DD model-based assistance, we observed that nP and Tk show the best task performance behavior when the user moves on the path (Fig. 6.1a). In addition, In assistance achieves the lowest task performance measure for "on path" movements; a possible tendency on In to perform close to Na and Ha may be considered (Table 6.1).

[1] For the on-path movements $p = 2.098 \times 10^{-14}$ and the off-path movements $p = 1.366 \times 10^{-14}$

[2] $p = 1.687 \times 10^{-6}$



(a) Samples for movements on the path. The nominal task case behavior.

(b) Samples for movements off the path. The umodeled situations case behavior.

Figure 6.1.: Boxplot for the RMS value of the tracking error position; samples are divided in movements on and off path. No differences between groups are represented by (**) with $p > 0.05$.

Off-path tracking error position

A Friedman test revealed a significant effect of assistances strategies on tracking error when human moves off the path ($\chi^2(4) = 35.533$, $p < 0.001$[3]). All post-hoc test were conducted using a Wilcoxon test with Holm-Bonferroni correction. Significant differences were found between Na and all other assistance strategies $\{nP, Tk, In, Ha\}$ ($p < 0.05, r = \{0.837, 0.778, 0.707, 0.554\}$). Also, differences were found between Ha and the DD-model based assistance conditions $\{nP, Tk, In\}$ ($p < 0.05, r = \{0.813, 0.601, 0.46\}$). Finally, a difference was found between In and nP ($p < 0.05, r = 0.625$). In contrast, no significant differences were found between Tk and nP ($r = 0.306$), nor between Tk and In ($r = 0.259$). The differences are highlighted in Table 6.2.

Table 6.2.: p-value (p) and effect size (r) for RMS vakue of tracking-position error when the user moves off the path.

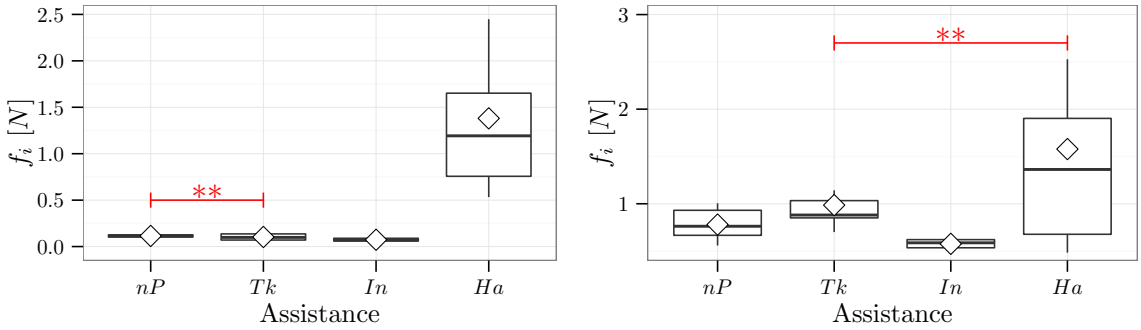
	nP		Tk		In		Ha	
	p	r	p	r	p	r	p	r
Tk	0.068	0.306	-	-	-	-	-	-
In	**0.005**	**0.625**	0.129	0.259	-	-	-	-
Ha	**0.005**	**0.813**	**0.005**	**0.601**	**0.037**	0.460	-	-
Na	**0.005**	**0.837**	**0.006**	**0.778**	**0.017**	**0.707**	**0.037**	**0.554**

From the previous statistical results, we can state that when the user moves off the path the tracking-position error is larger for Na compared to all other conditions, followed by Ha (Table 6.2 and Fig. 6.1b). nP and Tk tend to perform close on task performance when the user moves off the path ($p = 0.068$). Note that, In and Tk tend to peform close on task performance for "off path" events ($p = 0.129$). But, nP present lower tracking-position error compared with In.

6.1.2. Internal forces

We are mainly interested in internal forces when the user moves off the path. In this phase, the internal forces are increased because of the assistance exerts forces towards the nominal path; the

[3] $p = 3.609 \times 10^{-7}$



(a) Samples for movements on the path, the nominal task case.

(b) Samples for movements off the path, the unmodeled situation case.

Figure 6.2.: Boxplot for the internal forces. No differences between groups are represented by (**) with $p > 0.05$.

artificial assistance should reduce the internal forces in order to assume a collaborative strategy and avoids fights against the user desire, reducing the assistance level and consequently the internal forces. When the user moves on the path the user agrees the assistance, therefore we expected a low values for internal forces.

The normality of f_i cannot be established according to KS test[4] (neither, for on or off path movements).

On-path internal forces

A Friedman test revealed a significant effect of assistances conditions on internal forces (f_i) when the user moves on the path ($\chi^2(3) = 27.9$, $p < 0.001$[5]). All post-hoc test were conducted using a Wilcoxon test with Holm-Bonferroni correction. Significant differences among Ha and DD model-based assistance conditions are obtained $\{nP, Tk, In\}$ ($p < 0.01$, $r = \{0.849, 0.849, 0.849\}$), In and $\{nP,Tk\}$ ($p < 0.05$, $r = \{0.742, 0.365\}$). No significant difference was found between nP and Tk ($p > 0.05$, $r = 0.153$). See Table 6.3 and Fig. 6.2a.

Table 6.3.: p-value (p) and effect size (r) for internal forces (f_i) when the user moves on the path.

	nP		Tk		In	
	p	r	p	r	p	r
Tk	0.424	0.153	-	-	-	-
In	**0.003**	**0.742**	**0.024**	0.365	-	-
Ha	**0.003**	**0.849**	**0.003**	**0.849**	**0.003**	**0.849**

Off-path internal forces samples

A Friedman test revealed a significant effect of assistances strategies on internal forces (f_i) when the user moves off the path ($\chi^2(3) = 18.8$, $p < 0.001$[6]). All post-hoc test were conducted using a Wilcoxon test with Holm-Bonferroni correction. Significant differences were found among Ha

[4]For on-path movements $p = 1.48 \times 10^{-12}$ and for off-path movements $p = 8.882 \times 10^{-16}$

[5]$p = 3.812 \times 10^{-6}$

[6]$p = 0$

and $\{nP, In\}$ ($p < 0.05$, $r = \{0.377, 0.636\}$). In addition, differences among the DD model-based assistances were found, nP and $\{Tk, In\}$ ($p < 0.05$, $r = \{0.354, 0.625\}$), and Tk and In ($p < 0.01$, $r = 0.849$). No significant difference was found between Ha and Tk ($p > 0.05$, $r = 0.2$). See Table 6.4 and Fig. 6.2b.

Table 6.4.: p-value (p) and effect size (r) for internal forces (f_i) when the user moves off the path.

	nP		Tk		In	
	p	r	p	r	p	r
Tk	**0.037**	0.354	-	-	-	-
In	**0.034**	**0.625**	**0.003**	**0.849**	-	-
Ha	**0.037**	0.377	0.092	0.200	**0.034**	**0.636**

Turning now to the experimental evidence, we found that in both movements, on and off the path, the DD-model-based assistance reduced internal forces compared to Ha (Tables 6.3,6.4 and Fig. 6.2). The only difference was that in the off path movements, the internal forces tend to present a close behavior on Tk and Ha ($p = 0.092$, Table 6.4 and Fig. 6.2b).

Regarding the on-path movements amongst the DD model-based assistance, a tendency towards similar behavior was achieved in terms of agreement between nP and Tk. Furthermore, statistically the lowest internal forces was presented for In assistance (Table 6.3 and Fig. 6.2a).

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 6.4 and Fig. 6.2b).

6.1.3. Physical Dominance

We analyzed the physical dominance (PD_{12}) of the user over the assistance. This measure for Ha condition evaluates the subject dominance over the human-collaborator; while for the DD model-based assistance conditions this metric evaluates the subject dominance over the artificial assistance. We include an analysis of the physical dominance difference (PD_{diff}), which evaluates the dyadic dominance during the interaction.

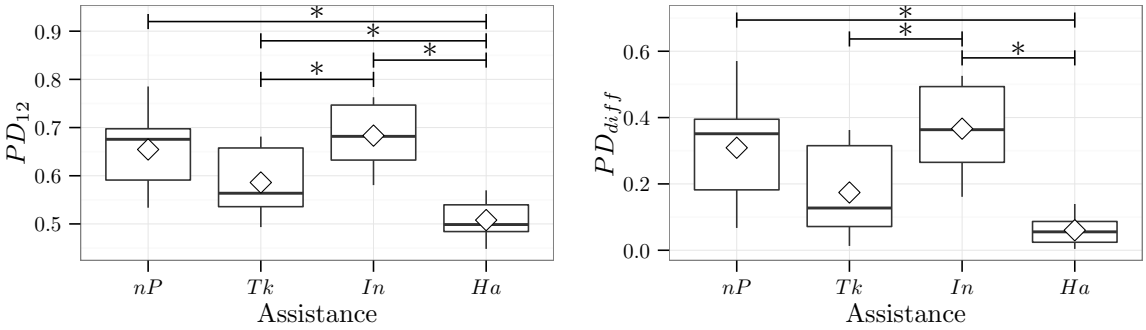
Physical dominance of the user over the assistance (PD_{12})

A Friedman test revealed a significant effect of assistance strategies on physical dominance from the subject over the assistance ($\chi^2(3) = 28.3$, $p < 0.001$[7]). Post-hoc test—all conducted using Wilcoxon test with Holm-Bonferroni correction—shows the significant differences between Ha and all other schemes $\{nP, Tk, In\}$ ($p < 0.05$, $r = \{0.778, 0.577, 0.849\}$) and between In and Tk ($p < 0.01$, $r = 0.577$). Differences are highlighted in Table 6.5 and Fig. 6.3a.

Table 6.5.: p-value (p) and effect size (r) for physical dominance from subject over the assistance (PD_{12}).

	nP		Tk		In	
	p	r	p	r	p	r
Tk	0.068	0.448	-	-	-	-
In	0.176	0.153	**0.003**	**0.577**	-	-
Ha	**0.003**	**0.778**	**0.010**	**0.577**	**0.003**	**0.849**

[7] $p = 3.142 \times 10^{-6}$



(a) PD_{12} (b) PD_{diff}

Figure 6.3.: Boxplot for the physical dominance of the user over the assistance (PD_{12}) and the physical dominance diference (PD_{diff}). The differences between groups are shown with a 0.05 level of significance (*).

Physical dominance difference (PD_{diff})

A Friedman test revealed a significant effect of assistance strategies on physical dominance difference ($\chi^2(3) = 26.7$, $p < 0.001$[8]). All post-hoc test were conducted using Wilcoxon test with Holm-Bonferroni correction, which showed the significant differences between Ha and $\{nP, In\}$ ($p < 0.01$, $r = \{0.754, 0.849\}$), and between In and Tk ($p < 0.01$, $r = 0.577$). In contrast, no significant differences were found between nP and $\{Tk, In\}$ ($p > 0.05, r = \{0.44, 0.17\}$), nor between Tk and Ha ($p < 0.05, r = 0.46$). See Table 6.6 and Fig. 6.3b.

Table 6.6.: 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	**0.003**	**0.577**	-	-
Ha	**0.003**	**0.754**	0.081	0.460	**0.003**	**0.849**

Overall, statistical evidence on physical dominance measures (PD_{12} and PD_{diff}) reveals differences between Ha and nP ($p = 0.003$), between Ha and In ($p = 0.003$), as well as between Tk and In ($p = 0.003$). Furthermore, on average a possible tendency towards similar behavior was found between nP and Tk, and between nP and In (Tables 6.5 and 6.6).

Interestingly, the unique difference between measures PD_{12} and PD_{diff} occurs for the comparison between Ha and Tk, in which significant difference was found for PD_{12} ($p = 0.01$), while no significant differences was found in PD_{diff} ($p = 0.08$). Furthermore, the effect size test shows a large effect in PD_{12} ($r = 0.57$) and medium effect in PD_{diff} ($r = 0.46$). This occurs between high assistance level conditions (Tk and Ha).

Regarding the experimental evidence, Ha presents the most sharing dominance strategy. While dominance in Tk presents most sharing strategy compared to In, Fig. 6.3.

[8] $p = 6.805 \times 10^{-6}$

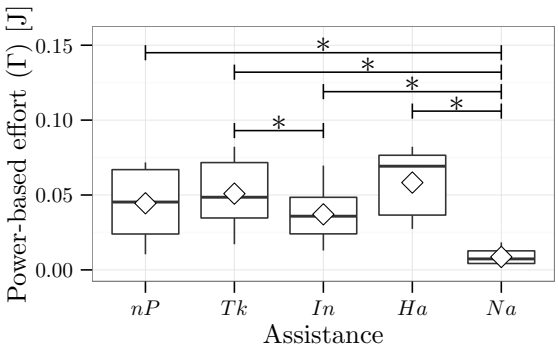


Figure 6.4.: Boxplot for the total power-based effort, the dyadic case (Γ). The differences between groups are shown with a 0.05 level of significance (*).

6.1.4. Power-based effort

This subsection reports physical effort measure calculated via the total force and the force exerted in the perpendicular direction of the desired path; for each case the individual and dyadic power is presented.

Total power-based effort measure, the dyadic case

A Friedman test revealed a significant effect of assistances strategies on power-based effort for the dyadic case, the total force exerted during the interaction is considered (Γ) ($\chi^2(3) = 26.8$, $p < 0.001$[9]). All post-hoc test were conducted using Wilcoxon test with Holm-Bonferroni correction, which showed the significant differences amongst Na and all schemes $\{nP, Tk, In, Ha\}$ ($p < 0.01$, $r = \{0.766, 0.837, 0.79, 0.849\}$). A difference was found between In and Tk ($p < 0.05$, $r = 0.318$). In contrast, no significant difference can be found amongst the other conditions for the DD model-based assistance, nor between Ha and the DD model-based assistances. Differences are highlighted in Table 6.7 and Fig. 6.4.

Table 6.7.: p-value (p) and effect size (r) for dyadic power-based effort (Γ).

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

Total power-based Effort, the individual case

The total effort measure for the individual case is defined as the mean absolute power (MAP) in which the total force exerted by the human is considered,

$$MAP^1 = \frac{1}{N}\sum_{\kappa=1}^{N}|P^1_{,\kappa}|$$

[9] $p = 6.484 \times 10^{-6}$

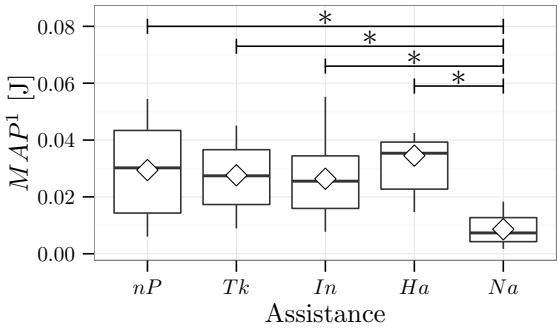


Figure 6.5.: Boxplot for the individual power-based effort (MAP^1). The effort due to the force exerted by the subject is shown. The differences between groups are shown with a 0.05 level of significance (*), Table 6.8.

where, the power ($P_{1,k}$) is calculated for the user (without considering the assistance power) in all experimental conditions.

As the normality cannot be assumed based on KS test[10], a Friedman test is conducted which revealed a significant effect of assistances strategies on individual power-based effort (the total force exerted by the human is considered) ($\chi^2(4) = 23.8$, $p < 0.001$[11]). All post-hoc test were conducted using Wilcoxon test with Holm-Bonferroni correction showed the significant differences between Na and all other conditions $\{nP, Tk, In, Ha\}$ ($p < 0.01$, $r = \{0.648, 0.731, 0.672, 0.79\}$). In contrast, no significant differences on effort measure were found amongst DD model-based assistance conditions; nor when DD model-based assistance are compared to Ha. See Fig. 6.5 and Table 6.8.

Table 6.8.: p-value (p) and effect size (r) for individual power-based effort (MAP^1)

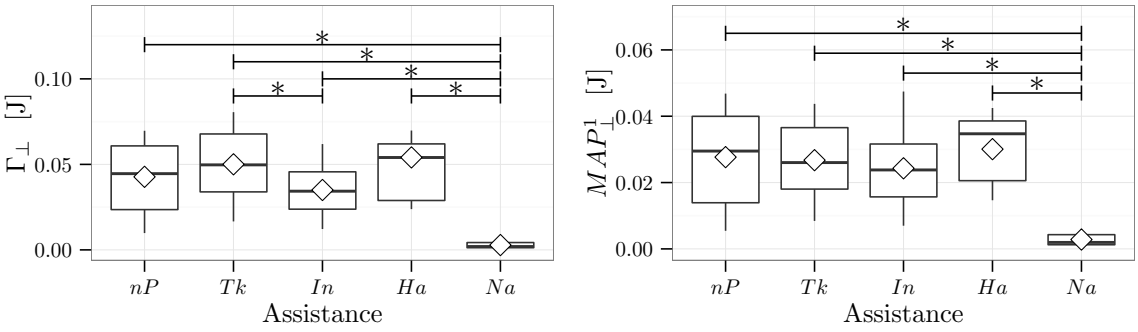
	nP		Tk		In		Ha	
	p	r	p	r	p	r	p	r
Tk	1.000	0.024	-	-	-	-	-	-
In	1.000	0.071	1.000	0.082	-	-	-	-
Ha	1.000	0.059	1.000	0.153	1.000	0.236	-	-
Na	**0.005**	**0.648**	**0.005**	**0.731**	**0.007**	**0.672**	**0.005**	**0.790**

Next, the statistical results for the physical effort in the perpendicular direction of the nominal path are presented. We include this analysis in order to explore the effort in the direction in which the assistance operates, for the dyadic and individual cases.

Perpendicular power-based effort, the dyadic case

We can not assume normality based on KS test[12]. Thus, a Friedman test revealed a significant effect of assistances schemes on the dyadic power-based effort which is calculated for the perpendicular direction of the path ($\chi^2(4) = 30.933$, $p < 0.001$[13]). A post-hoc test using Wilcoxon test with Holm-Bonferroni correction showed the significant differences between Na and all other conditions

[10] $p = 1.987 \times 10^{-14}$
[11] $p = 8.76 \times 10^{-5}$
[12] For individual case $p = 2.209 \times 10^{-14}$ and for the dyadic case $p = 2.209 \times 10^{-14}$
[13] $p = 3.159 \times 10^{-6}$



(a) Dyadic case

(b) Individual case

Figure 6.6.: Boxplot for the power-based effort calculated in the perpendicular direction of the nominal path. The dyadic (Γ) and the individual effort (MAP^1) are shown. The differences between groups are shown with a 0.05 level of significance (*).

$\{nP,\ Tk,\ In,\ Ha.\}$ ($p < 0.01$, $r = \{0.849, 0.849, 0.849, 0.849\}$). Also a difference was found between Tk and In ($p < 0.01$, $r = 0.342$). In contrast, no significant differences were found amongst DD model-based assistance conditions, neither when they are compared to Ha. See Table 6.9 and Fig. 6.6a.

Table 6.9.: p-value (p) and effect size (r) for the **dyadic** power-based effort in the **perpendicular direction** of the path ($\Gamma_\perp$).

	nP		Tk		In		Ha	
	p	r	p	r	p	r	p	r
Tk	0.212	0.177	-	-	-	-	-	-
In	0.454	0.165	**0.009**	0.342	-	-	-	-
Ha	0.939	0.189	0.939	0.071	0.256	0.389	-	-
Na	**0.005**	**0.849**	**0.005**	**0.849**	**0.005**	**0.849**	**0.005**	**0.849**

Perpendicular power-based effort, the individual case

A Friedman test revealed a significant effect of assistances schemes on individual power-based effort, which is calculated for the force exerted by the subject in the perpendicular direction of the nominal path ($\chi^2(4) = 26.867$, $p < 0.001$[14]). A post-hoc test using Wilcoxon test with Holm-Bonferroni correction showed the significant differences between Na and all other conditions $\{nP,\ Tsk,\ In,\ Ha\}$ ($p < 0.01$, $r = \{0.825, 0.849, 0.849, 0.849\}$). No significant differences among DD model-based assistance were obtained, neither when was compared to Ha. See Table 6.10 and Fig 6.6b.

Note that no significant statistical differences were found for the effort calculated based on total force and perpendicular force for both individual and dyadic cases. The dyadic case exhibits a difference between In and Tk, which cannot be assumed in the individual case (for both total and perpendicular 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 6.7 and Fig. 6.4)

[14] $p = 2.115 \times 10^{-5}$

Table 6.10.: p-value (p) and effect size (r) for the **individual** power-based effort in the **perpendicular direction** of the nominal path ($MAP^1_\perp$).

	nP		Tk		In		Ha	
	p	r	p	r	p	r	p	r
Tk	1.000	0.035	-	-	-	-	-	-
In	1.000	0.094	1.000	0.165	-	-	-	-
Ha	1.000	0.106	1.000	0.189	0.776	0.306	-	-
Na	**0.005**	**0.825**	**0.005**	**0.849**	**0.005**	**0.849**	**0.005**	**0.849**

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

6.1.5. Efficiency

The normality can be assumed based on the KS test for the efficiency measure[15]. Thus, a one-way, repeated-measurement ANOVA was conducted for efficiency measure.

Mauchly's test did not show a violation of sphericity against assistance conditions for efficiency ($W(4) = 0.252, p = 0.169$). With one way repeated measure ANOVA, we found a significant effect of assistance conditions on efficiency ($F(4, 44) = 11.126, p < 0.001$[16], partial $\eta^2 = 0.503$, $\eta^2 = 0.385$).

A pairwise comparison with the Holm-Bonferroni correction revealed the significant differences between Ha and all other conditions $\{nP,Tk,In,Na\}$ ($p < 0.05$); as well as between Na and nP ($p < 0.05$). Table 6.11 and Fig. 6.7a.

Table 6.11.: p-value for pairwise comparison for efficiency Λ.

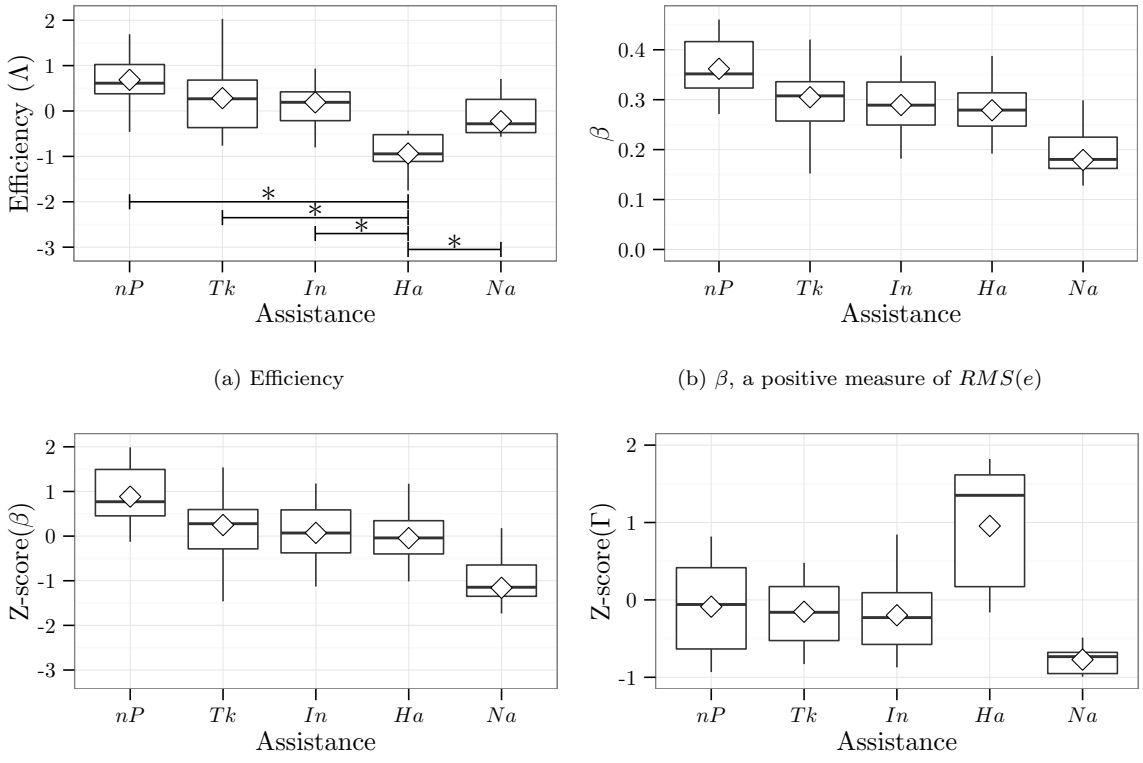
	nP	Tk	In	Ha
Tk	0.223	-	-	-
In	0.221	0.660	-	-
Ha	**0.009**	**0.029**	**0.028**	-
Na	**0.029**	0.223	0.260	**0.029**

In terms of the experimental evidence, Ha is the least efficient amongst the dyadic strategies and no-assistance condition. Amongst DD model-based assistance no significant differences on efficiency was found. Interestingly, no significant efficiency was found between DD model-based assistance and no-assistance (Na) condition, except for nP which improved efficiency compared to Na.

6.1.6. Physical dominance correlation with efficiency and task performance

We correlate physical dominance difference with efficiency (Fig. 6.8), which analyses whether the adopted physical dominance strategy may be interpreted as an efficiency strategy.

[15] $p = 0.638$
[16] $p = 2.537 \times 10^{-6}$



(a) Efficiency

(b) β, a positive measure of $RMS(e)$

(c) Performance - β, z-score

(d) Power-based effort - Γ, z-score

Figure 6.7.: Boxplot for the efficiency (Λ). The differences between groups are shown with 0.05 level of significance (*). Besides the z-score for β which represents the positive measure of $RMS(e)$, and effort measure (Γ) are shown.

Table 6.12.: p-value for Pearsons correlation test. Correlation for PD_{diff} with efficiency (Λ) is shown.

	nP	Tk	In	Ha
$PD_{diff} - \Lambda$	**0.004**	0.908	0.087	0.657

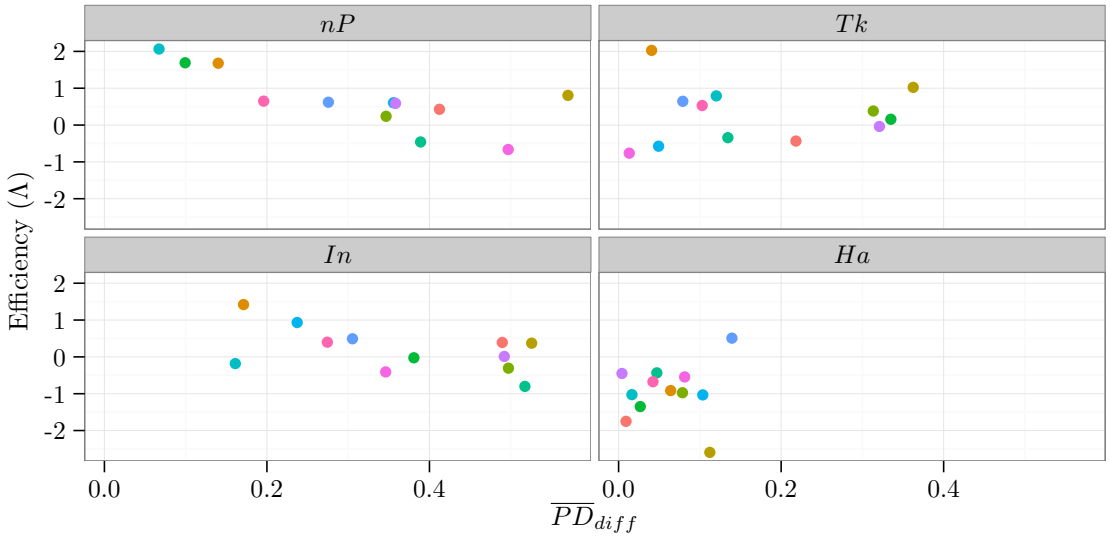


Figure 6.8.: Correlation between efficiency (Λ) and physical dominance difference ($\overline{PD}_{diff}$). Pearson test shows correlation for nP condition (r = 0.49).

Efficiency (Λ) presented no-correlation with physical dominance difference for Tk, In and Ha. In contrast, nP presented correlation between efficiency and PD_{diff} (pearson's $r(10) = 0.495, p < 0.01$).

6.2. Subjective perception results

Of the study population, 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 had previous experience with the interface.

Half of the operators perceived the strategy that leads to the best interaction performance to be the nP condition, followed by the Ha and In conditions (Fig. 6.9). Regarding the perceived overall task performance no major trend was found. Four operators perceived that the In condition allows to perform the task with minimum tracking error, while other three operators perceive that the best strategy is the Ha condition. nP and Tk conditions were chosen by two operators and just one user chooses Na.

Having presented the statistical results, we will now move on to discuss the hypothesis.

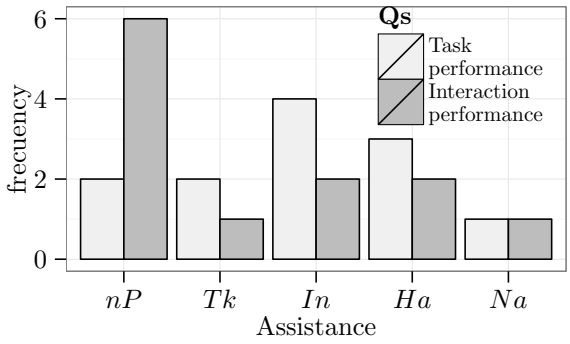


Figure 6.9.: Histogram of preferred conditions when judging interaction and task performance.

6.3. Discussion

The presented analysis considers the statistical results in the context of the main hypothesis. Thus, the results allow conclusions on the DD model-based haptic assistance dynamics and its comparison with human-like assistance and no-assistance strategies.

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

Regarding on-path movements (Fig. 6.1a), low assistance level in In cannot provide the required assistance level to improve nominal task performance compared with Na—actually, a possible tendency towards a similar behavior is obtained. In contrast, nP and Tk provide a high assistance level in nominal task execution. Thus task performance is improved compared with Na.

In telerobotics, Boessenkool et.al. [Boessenkool et al., 2013] reported similar results, haptic assistance improves task performance compared with no-assistance in a structured environment; in our study a similar result was obtained but extended to a partially structured environment.

Regarding off-path movements (Fig. 6.1b), a large tracking-position error may suggest a non-optimal strategy adopted to circumventing the obstacle, i.e. the user spends more time performing abrupt movements. We assumed the user wants to move closer to the path which was drawn to overcoming the obstacles (Fig. 5.1b). Internal forces supply a more convenient measure to establish performance in the off-path movements. In fact, it is internal forces rather than tracking-position error which was optimized by the DD model-based assistance when the user moves off the path.

Regarding off-path movements, the mean value of tracking-position error across all conditions is greater than the dimensions of the obstacle[17], this suggest that users, on average, avoid successfully the obstacles in all conditions, Fig. 6.10.

Thus, H1 can be partially confirmed. Despite nP and Tk improve task performance compared to Na in on-path situations, In behavior presents a possible tendency towards performing close to Na. Thus, we cannot assure that In improves task performance compared to Na in on-path situations. In off-path situations all DD-model-based assistances improve task performance with respect to Na, but at different rates with nP performing best.

[17]The perpendicular dimension of the obstacle is 0.5 cm. On average, the error when the user moves off path for each condition is grater than 0.9 cm, see Figs. 5.1b and 6.10

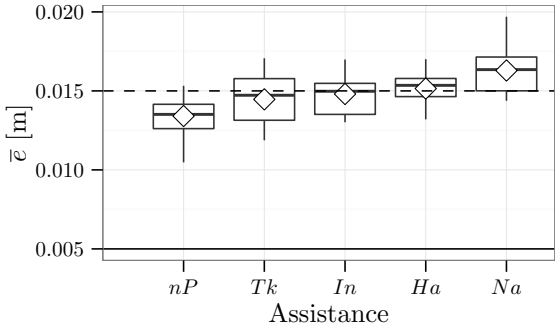


Figure 6.10.: Boxplot for the mean value of tracking-position error ($\overline{e}$). Only the off-path samples suggested to overcome the obstacle—parallel to the desired path—are shown. The solid line along 0.005 shows the obstacle dimension, while the dotted line along 0.015 shows the dimension for the suggested path to overcome the obstacles, cf. Fig. 5.1b.

H2. DD model-based assistance performs as a human for task performance, interaction performance and physical effort

Regarding tracking-position error for on-path movements, In performs close to Ha (Fig. 6.1a), while nP and Tk 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.

All DD model-based assistances successfully reduced internal forces compared to Ha in the on-path condition (Fig. 6.2a). While Tk tends to perform close on average to Ha for off-path conditions, nP and In successfully reduced internal forces compared with Ha (Fig. 6.2b)

In general, physical dominance is shared between human and assistance in all conditions—the PD_{diff} value is less than 0.5, Fig. 6.3. Nevertheless, Ha is clearly the condition with the most equal physical dominance strategy. 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 human-assistant. Furthermore, high assistance levels in the Tk condition, exhibited a possible tendency to perform close to the human-like condition (Ha) respecting the physical dominance degree. 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 actively the movements on the path.

Interestingly, in energy terms the user may perceive on average similar effort when executes the task with a human collaborator than with the DD model-based assistance. DD model-based did not exceed the physical effort values obtained in the human-like condition (Ha), Figs 6.4 and 6.5.

Overall, *H2 cannot be confirmed*. Significant differences between Ha and DD-model-based assistances were found for task performance, interaction performance, and dominance. The Tk condition performed closest to the Ha condition in terms of interaction performance and physical dominance, while the In condition performed closest for task performance. Collaborating humans in the Ha condition achieved an overall more balanced physical dominance, used higher interaction forces and achieved a worse task performance compared to the realized DD-model-based assistances. Only the overall effort of Ha and DD model-based assistances shows a tendency towards similar behavior.

Summarizing, the realized DD model-based assistances seem to be more compliant than a human collaborator as they leave more physical dominance with the user and thus, allow achieving an overall better task performance. Efforts spent 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.

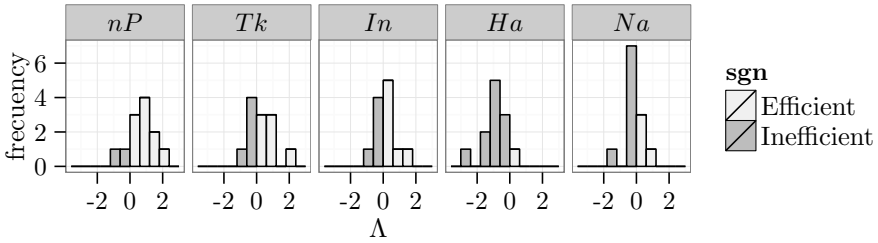


Figure 6.11.: Histogram for efficiency in the assistance conditions. Red bars indicates the number of operators that efficiently executes the task.

This provides evidence that a slightly unequal dominance distribution as realized in the DD-model-based assistances may benefit task performance. It should be noted though that such unequal dominance distribution could be also realized for the Ha condition if proper instructions were given, which may have led to a higher agreement of results achieved for the Ha and DD-model-based assistance conditions.

H3. DD model-based assistance will improve user efficiency compared with Na and Ha

Efficiency must be interpreted with caution because of DD model-based assistance and Ha conditions are dyadic schemes, while Na consist in an individual scheme in which no additional effort due to interaction forces is presented.

Statistical evidence reveals that efficiency (Λ) is improved when the human executes the task with the DD model-based assistance compared to Ha. Actually the number of subjects performing the task efficiently is larger when human execute the task with the DD model-assistance compared with human assistance (Ha), Fig. 6.11. The Ha exhibits the minimum number of subjects that efficiently perform the task (1/12) followed by the Na (4/12). This might suggest that subjects adapt more naturally with nP assistance in order to efficiently execute the task. The subjective perception confirm this trend (cf. Sec. 6.3.1).

Comparing the efficiency achieved by Na with Tk and In, a possible tendency towards performing close on average efficiency was found. While efficiency is improved for nP compared with Na ($p = 0.029$).

The number of subjects that performs efficiently the task with a nP is grater (10/12) than the exposed by the Tk (7/12) or In (7/12), Fig. 6.11. However, the post-hoc test showed that efficiency amongst the DD model-based assistance were not statistically significant (Table 6.11).

We may see Ha and Na as extreme cases on interaction. Ha requires additional effort in order to improve task performance, while Na achieves the worst task performance with no-additional effort. Statistical difference and the number of subjects executing efficiently the task support this trend.

Note that no-correlation between physical dominance and efficiency was found—only in the nP scheme was found. We expected that more shared physical dominance strategies relates with better efficiency outcomes because effort is shared between the dyad. Actually, as is shown in Fig. 6.8 and 6.11, nP assistance seems to improve efficiency when a shared dominance strategy is adopted. Nevertheless, further research is necessary to explore the relation between optimal decision-making strategies based on the matching shoulder parameters and the achieved efficiency, a possible approach is proposed in congnitive studies (see e.g. [Bogacz et al., 2006]).

H3 can be partially confirmed. Despite DD model-based assistance improve efficiency compared

to Ha, DD model-based assistance tends to performe close to Na regarding efficiency measure (note that Na consist in an individual strategy, in which no additional effort due to interaction is considered). We may suggest that DD model-based assistance take advantages of both extreme scenarios, Na and Ha.

H4. The DD model-based assistance reproduces different assistance behaviors by means of the reward structure setup

We investigated the influence of the reward structure on the haptic assistance (H4) by analysing the different dynamics obtained for task performance, physical efforts and interaction performances namely, internal forces and physical dominance difference.

Regarding tracking-position error in on-path movements, Tk and nP set the best task performance while In tends to perform close to Ha and Na. The configuration of reward functions for In provides low assistance level and enabled the user to adopt free movements reproducing a close response of the Ha or Na.

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 s/he 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 (Tables 6.3,6.4 and Fig. 6.2), 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 assistance, 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. 6.3 and Table 6.6), while nP can be found in-between. This suggests that the reward functions successfully allocated the physical dominance in the DD model-based assistances.

Regarding physical effort, In condition 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 in-between.

Thus, overall *we can confirm H3* 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.

6.3.1. User perception

The qualitative questionnaire provides some insight into the subjectively perceived task and interaction performances, Fig. 6.9. Although there is no clear trend, subjects perceive an improvement in terms of task performance with respect to the not assisted case (Na). Further, users tend to say that better task performance is achieved with low assistance levels (In) and the assistance of a human expert (Ha), which, however does not agree with objective tracking performance results.

Quantitatively, the best tracking performance is achieved under the nP condition. This condition was rated as the most preferred one in terms of interaction performance. Only one user found the Tk condition with the highest levels of assistance as well as the condition without any assistance (Na) as best in terms of interaction performance. This indicates that a high assistance

level or no assistance have a negative effect on the perceived difficulty level when performing the given task as either the user has to fight against the assistance or is not supported at all.

Four operators perceive that the In condition allows to perform the task with minimum tracking error, while other three operators perceive that the best strategy is the Ha condition. nP and Tk conditions were chosen by two operators and just one user chose Na.

This last preference has been reported in other studies where certain subjects consider any assistance as invasive and prefer to have total control on task [Powell and O'Malley, 2011].

7. Conclusions

Affine system stability analysis has been performed on the parameter-varying control architecture in order to derive theoretical result for the adoption of a varying assistance scheme. As a result, sufficient stability conditions have been provided such that under any bounded variation of the stiffness the system remains stable. This allows the designer to develop suitable scheduling strategies which improve performance metrics such as task performance (tracking in this case) or/and interaction performance.

A novel haptic assistance scheme has been proposed, such that a decision-making mechanism chooses between enhancing tracking performance or human-machine interaction via the scheduling of admittance control parameters; the proposed assistance successfully solves the trade-off between assisting the user's motion (which aims at improving task performance) and providing the user with enough freedom (in order to be in charge of the task when unmodeled situations occur in the nominal task). Two decision making models, the *Drift-Diffusion Model* and the γ-*policy*, were individually incorporated into the haptic assistance mechanism as the specific sharing strategies.

The results showed that the DD model is a suitable scheduling strategy for haptic shared control or other control allocation tasks as demonstrated by experimental data. The decision making process can be influenced via the parameters of the reward functions. On the contrary, the γ-policy exhibits a chattering behavior in the decision-making process which produces oscillations in the provided assistance level (the stiffness), degrading user experience. The main drawback in the γ-policy for haptic assistance is that it just considers the last two rewards. It has been observed that including more "memory" in the haptic assistance, as occurs in human decision making processes, provides the assistance with better suited performance.

All the experimental conditions related with the DD model-based assistance improve performance over the no assisted condition. Furthermore, the parameters of the reward structure determine the decision making process, which permit in turn to adapt the assistance behavior to different task requirements. For instance, Tk improves task performance, but at the same time it increases internal forces. In reduces internal forces but also reduces task performance because of the low assistance level is provided. nP seems to adapt better to user behavior due to the fact that both criteria—task and interaction measures—were balanced. This trend is also observed via the efficiency results analysis and the subjective perception of the users.

The effort increases in Tk compared to In but both assistance schemes present a similar improvements compared to Ha. The physical dominance of task execution is also affected by the reward structure. Ha is the most shared physical dominance condition, while for In the assistance gives the control to the user due to the provided low assistance. Significant differences between the performance of the DD model-based assistance and the performance of human-human collaboration (operating at equal physical dominance levels) were observed, indicating that the dominance distribution plays an important role and an unequal distribution may benefit overall task performance.

Limitations and recommendations

Although path following and obstacle avoidance tasks are present in several applications, the proposed assistance has to be adjusted to the particular application. Thus, the assistance design involves experimental procedures in which the system parameters are determined.

We explored the simplest decision-making strategy, i.e. choosing between two alternatives. Since

the haptic feedback encodes these alternatives into the haptic assistance, the human understands the mechanism and acts cooperatively. Further research to apply this concept to a multiple alternative choice scenarios is necessary. A promising approach may be found in cognitive science (see e.g. [Nedic et al., 2008, Krajbich and Rangel, 2011]), in which extensions for the drift-diffusion model to a multiple choice are proposed.

Although dynamic decision making is a promising allocation strategy in human-robot applications, further research is still necessary in order to determine what a "natural" assistance policy means and how can it be achieved. This work has demonstrated that bio-inspired mechanisms lead to a suitable solutions, specifically decision-making models.

Future directions

Future work may consider proposing a formal algorithm in order to configure the different parameters of the DD model-based assistance. Possible approaches can be found in the cognitive literature (e.g. [Bogacz et al., 2006]). Future work will also include:

- Adapting the level of assistance based on the adaptation of the slope of the matching shoulder structure. This varying assistance may depend on new incoming information of the task.
- Scheduling between more than two levels by including voting mechanisms
- Exploring alternative decision-making models which may include the decision making processes of the partner into the assistance mechanism

Appendices

A. Computed Torque

The robot manipulator dynamics can be described (in the robot's workspace) as [Craig, 2005]:

$$\mathbf{M}_x(\mathbf{q})\ddot{\mathbf{x}} + \mathbf{C}_x(\mathbf{q}, \ddot{\mathbf{q}})\dot{\mathbf{x}} = \mathbf{f}_a + \mathbf{f}_{ext} \tag{A.1}$$

where $\mathbf{M}_x$ is a positive definite matrix of mass (inertia), $\mathbf{C}_x$ is a vector containing Coriolis, centrifugal and gravity forces, $\mathbf{f}_a$ is the effective force exerted by the actuators on the end effector, $\mathbf{f}_{ext}$ is an external force applied to the end effector, $\mathbf{q}$ is the generalized vector of joint variables and $\mathbf{x}$ are the generalized coordinates in the robot's workspace coordinates.

$\mathbf{f}_a$ is expressed in the joint space, using the Jacobian matrix $\mathbf{J}$,

$$\boldsymbol{\tau}_a = \mathbf{J}^T(\mathbf{q})\,\mathbf{f}_a \tag{A.2}$$

where $\boldsymbol{\tau}_a$ are the forces (torques) exerted by the actuators and matrices $\mathbf{M}_x$ and $\mathbf{C}_x$ are obtained from the description in the joint space of $\mathbf{M}$ and $\mathbf{C}$, respectively [Craig, 2005]:

$$\mathbf{M}_x(\mathbf{q}) = \mathbf{J}^T\,\mathbf{M}(\mathbf{q})\,\mathbf{J}^{-1} \tag{A.3}$$

$$\mathbf{C}_x(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{J}^{-T}\left(\mathbf{C}(\mathbf{q}, \ddot{\mathbf{q}}) - \mathbf{M}(\mathbf{q})\mathbf{J}^{-1}\,\dot{\mathbf{J}}\,\dot{\mathbf{q}}\right) \tag{A.4}$$

Where, the Jacobian ($\mathbf{J}$) depends on the generalized coordinates in the joint space, $\mathbf{J}(\mathbf{q})$.

This model requires the calculation of the inverse of the Jacobian; i.e the Jacobian must be square and nonsingular.

Considering the computed torque control technique [Craig, 2005] (also called by inverse dynamics [Spong et al., 2005]), the nonlinear dynamics of the robot (A.1) can be linearizing and decoupled, i.e. the new controlled dynamics of the system are represented by a double integrator,

$$\mathbf{f}_a = \cancel{m}\,\mathbf{1}\ddot{\mathbf{x}}$$

To achieve a full servo controller, the control is divided into two parts: one is focused on a model based on a system linearisation, while concentrates on the other to reducing errors in the model (servo part) with prescribed dynamics.

Consider the control law [Craig, 2005], in which the effective force exerted on the end effector of the robot is calculated by,

$$\mathbf{f}_a = \boldsymbol{\nu}\mathbf{f}'_a + \boldsymbol{\iota}$$

where, $\boldsymbol{\nu}$ and $\boldsymbol{\iota}$ are given by,

$$\boldsymbol{\nu} = \mathbf{M}_x(\mathbf{q})$$
$$\boldsymbol{\iota} = \mathbf{C}_x + \mathbf{f}_{ext}$$

Now consider a feedforward law of the desired acceleration is included. The desired acceleration ($\ddot{\mathbf{x}}$) is calculated on or off-line, where,

$$\mathbf{f}'_a = \ddot{\mathbf{x}}_d + \mathbf{K}_v\,\dot{\mathbf{E}} + \mathbf{K_p}\,\mathbf{E} \tag{A.5}$$

The closed-loop system is then characterized by the error equation:

$$\ddot{\mathbf{E}} + \mathbf{K}_v\dot{\mathbf{E}} + \mathbf{K}_p\mathbf{E} = \mathbf{0}$$

where, $\mathbf{K}_v$ and $\mathbf{K}_p$ are diagonal matrices and it is desired that each joint present a critically damped behavior.

Since in practice it is difficult to ensure accurate models of the matrices $\mathbf{M}_x$, $\mathbf{C}_x$ and $\mathbf{f}_{ext}$, robust control and adaptive strategies have been proposed [Craig, 2005,Spong et al., 2005] to address plan model uncertainties.

Bibliography

[3D-Systems,] 3D-Systems. Geomagic touch (formerly geomagic phantom omni) overview. http://www.geomagic.com/en/products/phantom-omni/overview/. Accessed: 2016-03-18.

[Aarno et al., 2005] Aarno, D., Ekvall, S., and Kragic, D. (2005). Adaptive virtual fixtures for machine-assisted teleoperation tasks. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation, 2005. ICRA 2005*, pages 897–903.

[Abbink et al., 2012] Abbink, D. A., Mulder, M., and Boer, E. R. (2012). Haptic shared control: smoothly shifting control authority? *Cogn. Technol. Work*, 14(1):19–28.

[Abbott et al., 2005] Abbott, J. J., Marayong, P., and Okamura, A. M. (2005). Haptic virtual fixtures for robot-assisted manipulation. In *12th International Symposium of Robotics Research (ISRR)*, pages 49–64.

[Abbott and Okamura, 2006] Abbott, J. J. and Okamura, A. M. (2006). Pseudo-admittance bilateral telemanipulation with guidance virtual fixtures. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2006 14th Symposium on*, pages 169–175. IEEE.

[Adams and Hannaford, 1999] Adams, R. and Hannaford, B. (1999). Stable haptic interaction with virtual environments. *IEEE Transactions on Robotics and Automation*, 15(3):465–474.

[Anderson and Spong, 1989a] Anderson, R. and Spong, M. (1989a). Asymptotic stability for force reflecting teleoperators with time delays. In *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on*, pages 1618–1625. IEEE Comput. Soc. Press.

[Anderson and Spong, 1989b] Anderson, R. and Spong, M. (1989b). Bilateral control of teleoperators with time delay. *IEEE Transactions on Automatic Control*, 34(5):494–501.

[Azhmyakov et al., 2011] Azhmyakov, V., Egerstedt, M., Fridman, L., and Poznyak, A. (2011). Approximability of nonlinear affine control systems. 5(2):275–288.

[Azorin et al., 2001] Azorin, J., Aracil, R., Sabater, J., and Fernandez, C. (2001). Bilateral control of teleoperators with time-delay: stability analisys.

[Bao and L. Lee, 2007] Bao, J. and L. Lee, P. (2007). *Process Control - The Passive Systems Approach*. Springer-Verlag London.

[Bettini et al., 2004] Bettini, A., Marayong, P., Lang, S., Okamura, A., and Hager, G. (2004). Vision-assisted control for manipulation using virtual fixtures. *IEEE Transactions on Robotics*, 20(6):953–966.

[Bodnaruk and Simonov, 2015] Bodnaruk, A. and Simonov, A. (2015). Do financial experts make better investment decisions? *Journal of Financial Intermediation*, 24(4):514–536.

[Boessenkool et al., 2013] Boessenkool, H., Abbink, D., Heemskerk, C., van der Helm, F., and Wildenbeest, J. (2013). A task-specific analysis of the benefit of haptic shared control during telemanipulation. *IEEE Transactions on Haptics*, 6(1):2–12.

[Boessenkool et al., 2011] Boessenkool, H., Abbink, D., Heemskerk, C. J. M., and Van der Helm, F. C. T. (2011). Haptic shared control improves tele-operated task performance towards performance in direct control. In *World Haptics Conference (WHC), 2011 IEEE*, pages 433–438.

[Bogacz et al., 2006] Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review*, 113(4):700–765.

[Bogacz et al., 2007] Bogacz, R., McClure, S. M., Li, J., Cohen, J. D., and Montague, P. R. (2007). Short-term memory traces for action bias in human reinforcement learning. *Brain Research*, 1153:111–121.

[Boyd et al., 1994] Boyd, S., Ghaoui, L., Feron, E., and Balakrishnan, V. (1994). *Linear Matrix Inequalities in System and Control Theory*, volume 1. SIAM.

[Boyd and Vandenberghe, 2004] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, New York, NY, USA.

[Brogliato et al., 2010] Brogliato, B., Lozano, R., Maschke, B., and Egeland, O. (2010). *Dissipative Systems Analysis and Control: Theory and Applications*. Springer.

[Bruemmer et al., 2005] Bruemmer, D., Few, D., Boring, R., Marble, J., Walton, M., and Nielsen, C. (2005). Shared understanding for collaborative control. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 35(4):494–504.

[Busemeyer and Johnson, 2004] Busemeyer, J. R. and Johnson, J. G. (2004). Computational models of decision making. *Blackwell handbook of judgment and decision making*, pages 133–154.

[Cao et al., 2008] Cao, M., Stewart, A., and Leonard, N. E. (2008). Integrating human and robot decision-making dynamics with feedback: Models and convergence analysis. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 1127 –1132.

[Cao et al., 2010] Cao, M., Stewart, A., and Leonard, N. E. (2010). Convergence in human decision-making dynamics. *Systems & Control Letters*, 59(2):87–97.

[Chopra and Spong, 2004] Chopra, N. and Spong, M. (2004). Adaptive coordination control of bilateral teleoperators with time delay. In *43rd IEEE Conference on Decision and Control, 2004. CDC*, volume 5, pages 4540–4547 Vol.5.

[Christiansson et al., 2006] Christiansson, Mulder, and van der Helm (2006). Slave device stiffness and teleoperator stability. In *EuroHaptics Conference, Paris, France*.

[Cohen, 1988] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 2 edition edition.

[Colgate, 1993] Colgate, J. (1993). Robust impedance shaping telemanipulation. *IEEE Transactions on Robotics and Automation*, 9(4):374–384.

[Colgate and Hogan, 1988] Colgate, J. E. and Hogan, N. (1988). Robust control of dynamically interacting systems. *Int. Jour. Cont.*, 48:65–88.

[Coskunoglu and Weber, 1989] Coskunoglu, O. and Weber, E. U. (1989). A synthesis of descriptive and prescriptive models for decision making. In Jackson, M. C., Keys, P., and Cropper, S. A., editors, *Operational Research and the Social Sciences*, pages 653–660. Springer US. DOI: 10.1007/978-1-4613-0789-1_99.

[Craig, 2005] Craig, J. (2005). *Introduction to Robotics: Mechanics and Control*. Addison-Wesley series in electrical and computer engineering: control engineering. Pearson Education, Incorporated, 3 edition.

[Daniel and McAree, 1998] Daniel, R. W. and McAree, P. R. (1998). Fundamental limits of performance for force reflecting teleoperation. *The International Journal of Robotics Research*, 17(8):811–830.

[Desai and Yanco, 2005] Desai, M. and Yanco, H. (2005). Blending human and robot inputs for sliding scale autonomy. In *IEEE International Workshop on Robot and Human Interactive Communication, 2005. ROMAN 2005*, pages 537–542.

[Dias et al., 2008] Dias, M. B., Kannan, B., Browning, B., Jones, E. G., Argall, B., Dias, M. F., Zinck, M., Veloso, M. M., and Stentz, A. J. (2008). Sliding autonomy for peer-to-peer human-robot teams 1. In *, 10th International Conference on Intelligent Autonomous Systems (IAS 2008)*.

[Dragan and Srinivasa, 2013] Dragan, A. D. and Srinivasa, S. S. (2013). A Policy Blending Formalism for Shared Control. *The International Journal of Robotics Research*, 32.

[Egelman et al., 1998] Egelman, D. M., Person, C., and Montague, P. R. (1998). A computational role for dopamine delivery in human decision-making. *J. Cogn. Neurosci*, page 623–630.

[Enes and Book, 2010] Enes, A. and Book, W. (2010). Blended Shared Control of Zermelo's navigation problem. In *American Control Conference (ACC), 2010*, pages 4307–4312.

[Evrard and Kheddar, 2009a] Evrard, P. and Kheddar, A. (2009a). Homotopy-based controller for physical human-robot interaction. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 1 –6.

[Evrard and Kheddar, 2009b] Evrard, P. and Kheddar, A. (2009b). 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*, pages 45 –50.

[Ferre et al., 2007] Ferre, M., Buss, M., Aracil, R., Melchiorri, C., and Balaguer, C. (2007). *Advances in Telerobotics*, volume 31. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Feth et al., 2009a] Feth, D., Groten, R., Peer, A., and Buss, M. (2009a). Control-theoretic model of haptic human-human interaction in a pursuit tracking task. In *The 18th IEEE International Symposium on Robot and Human Interactive Communication, 2009. RO-MAN 2009*, pages 1106–1111. IEEE.

[Feth et al., 2009b] Feth, D., Tran, B. A., Groten, R., Peer, A., and Buss, M. (2009b). *Human Centered Robot Systems: Cognition, Interaction, Technology*, chapter Shared-Control Paradigms in Multi-Operator-Single-Robot Teleoperation, pages 53–62. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Fischer et al., 2015] Fischer, A., Holt, D., and Funke, J. (2015). Promoting the growing field of dynamic decision making. *Journal of dynamic decision making*, 1(1).

[Fite et al., 2004] Fite, K., Shao, L., and Goldfarb, M. (2004). Loop shaping for transparency and stability robustness in bilateral telemanipulation. *Robotics and Automation, IEEE Transactions on*, 20(3):620–624.

[Flash and Hogans, 1985] Flash, T. and Hogans, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *Journal of Neuroscience*, 5:1688—1703.

[Force-dimension,] Force-dimension. Force dimension - products - sigma.7 - overview. http://www.forcedimension.com/products/sigma-7/overview. Accessed: 2016-03-19.

[Furht, 2006] Furht, B. (2006). Haptics. In *Encyclopedia of Multimedia*, pages 266–267. Springer US.

[Gil et al., 2004] Gil, J., Avello, A., Rubio, A., and Florez, J. (2004). Stability analysis of a 1 DOF haptic interface using the Routh&Hurwitz criterion. *IEEE Transactions on Control Systems Technology*, 12(4):583–588.

[Gonzalez, 2013] Gonzalez, C. (2013). Cognitive science: An introduction. In Lanzer, P., editor, *Catheter-Based Cardiovascular Interventions*, pages 61–67. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-27676-7_6.

[Gonzalez and Quesada, 2003] Gonzalez, C. and Quesada, J. (2003). Learning in dynamic decision making: The recognition process. *Computational & Mathematical Organization Theory*, 9(4):287–304.

[Groten, 2011] Groten, R. (2011). *Haptic Human-Robot Collaboration: How to Learn from Human Dyads*. PhD thesis, Technische Universität München.

[Gunn et al., 2009] Gunn, C., Muller, W., and Datta, A. (2009). Performance improvement with haptic assistance: A quantitative assessment. In *EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*, pages 511–516.

[Haddad and Chellaboina, 2008] Haddad, W. M. and Chellaboina, V. (2008). *Nonlinear dynamical systems and control: a Lyapunov-based approach*. Princeton University Press, Princeton.

[Haddadi and Hashtrudi-Zaad, 2010] Haddadi, A. and Hashtrudi-Zaad, K. (2010). Bounded-impedance absolute stability of bilateral teleoperation control systems. *IEEE Transactions on Haptics*, 3(1):15–27.

[Hannaford, 1989] Hannaford, B. (1989). A design framework for teleoperators with kinesthetic feedback. *IEEE Transactions on Robotics and Automation*, 5(4):426–434.

[Hannaford et al., 1991] Hannaford, B., Wood, L., McAffee, D., and Zak, H. (1991). Performance evaluation of a six-axis generalized force-reflecting teleoperator. *Systems, Man and Cybernetics, IEEE Transactions on*, 21(3):620–633.

[Hashtrudi-Zaad and Salcudean, 2001] Hashtrudi-Zaad, K. and Salcudean, S. E. (2001). Analysis of control architectures for teleoperation systems with Impedance/Admittance master and slave manipulators. *The International Journal of Robotics Research*, 20(6):419–445.

[Hill et al., 2005] Hill, J. C., Johnson, F. R., Archibald, J. K., Frost, R. L., and Stirling, W. C. (2005). A cooperative multi-agent approach to free flight. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '05, pages 1083–1090. ACM.

[Hogan, 1989] Hogan, N. (1989). Controlling impedance at the man/machine interface. In *, 1989 IEEE International Conference on Robotics and Automation, 1989. Proceedings*, pages 1626–1631 vol.3.

[Hokayem and Spong, 2006] Hokayem, P. F. and Spong, M. W. (2006). Bilateral teleoperation: An historical survey. *Automatica*, 42(12):2035–2057.

[Huang and Chang, 2004] Huang, Y. and Chang, S. (2004). Table look-up variable structure control for robotic arms. In *Control Conference, 2004. 5th Asian*, volume 1, pages 585–591 Vol.1.

[Hulin et al., 2008] Hulin, T., Preusche, C., and Hirzinger, G. (2008). Stability boundary for haptic rendering: Influence of human operator. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3483–3488. IEEE.

[Johnson and Busemeyer, 2006] Johnson, J. G. and Busemeyer, J. R. (2006). A unified computational modeling approach to decision making. In *Proceedings of the Seventh International Conference on Cognitive Modeling*.

[Kelly et al., 1989] Kelly, R., Carelli, R., Amestegui, M., and Ortega, R. (1989). On adaptive impedance control of robot manipulators. In *, 1989 IEEE International Conference on Robotics and Automation, 1989. Proceedings*, pages 572–577 vol.1.

[Khalil, 2001] Khalil, H. K. (2001). *Nonlinear Systems*. Prentice Hall, 3 edition.

[Kim and Ahn, 2011] Kim, B.-Y. and Ahn, H.-S. (2011). Necessary and sufficient condition for the existence of a common quadratic lyapunov function for switched linear interval systems. In *2011 11th International Conference on Control, Automation and Systems (ICCAS)*, pages 97–102.

[Kragic et al., 2005] Kragic, D., Marayong, P., Li, M., Okamura, A. M., and Hager, G. D. (2005). Human-machine collaborative systems for microsurgical applications. *International Journal of Robotics Research*, 24:731–741.

[Krajbich and Rangel, 2011] Krajbich, I. and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33):13852–13857. PMID: 21808009.

[Kron and Schmidt, 2005] Kron, A. and Schmidt, G. (2005). Haptic telepresent control technology applied to disposal of explosive ordnances: Principles and experimental results. In *Industrial Electronics, 2005. ISIE 2005. Proceedings of the IEEE International Symposium on*, volume 4, pages 1505 – 1510.

[Kron et al., 1999] Kron, A., Schmidt, G., Baier, H., Buss, M., Freyberger, F., Hoogen, J., Kammermeier, P., and Schmidt, G. (1999). Distributed PC-based haptic, visual and acoustic telepresence system-experiments in virtual and remote environments. In *Virtual Reality, 1999. Proceedings., IEEE*, volume 4, pages 118 –125.

[Kronander and Billard, 2014] Kronander, K. and Billard, A. (2014). Learning Compliant Manipulation through Kinesthetic and Tactile Human-Robot Interaction. *IEEE Transactions on Haptics*, 7(3):367–380.

[Kucukyilmaz et al., 2013] Kucukyilmaz, A., Sezgin, T., and Basdogan, C. (2013). Intention Recognition for Dynamic Role Exchange in Haptic Collaboration. *IEEE Transactions on Haptics*, 6(1):58–68.

[Lawrence, 1993] Lawrence, D. (1993). Stability and transparency in bilateral teleoperation. *IEEE Transactions on Robotics and Automation*, 9(5):624–637.

[Li et al., 2007] Li, M., Kapoor, A., and Taylor, R. H. (2007). Telerobotic control by virtual fixtures for surgical applications. In Ferre, M., Buss, M., Aracil, R., Melchiorri, C., and Balaguer, C., editors, *Advances in Telerobotics*, number 31 in Springer Tracts in Advanced Robotics, pages 381–401. Springer Berlin Heidelberg.

[Li and Okamura, 2003] Li, M. and Okamura, A. M. (2003). Recognition of operator motions for real-time assistance using virtual fixtures. In *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2003. HAPTICS 2003. Proceedings. 11th Symposium on*, pages 125 – 131.

[Li et al., 2009] Li, Y., Huegel, J., Patoglu, V., and O'Malley, M. (2009). Progressive shared control for training in virtual environments. In *EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*, pages 332–337.

[Li et al., 2015] Li, Y., Tee, K. P., Chan, W. L., Yan, R., Chua, Y., and Limbu, D. K. (2015). Role adaptation of human and robot in collaborative tasks. In *2015 IEEE International Conference on Robotics and Automation (ICRA2015)*.

[Lin and Antsaklis, 2009] Lin, H. and Antsaklis, P. J. (2009). Stability and stabilizability of switched linear systems: A survey of recent results. *IEEE Transactions on Automatic Control*, 54(2):308–322.

[Marayong et al., 2002] Marayong, P., Bettini, A., and Okamura, A. (2002). Effect of virtual fixture compliance on human-machine cooperative manipulation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2002*, volume 2, pages 1089–1095 vol.2.

[Marayong et al., 2003] Marayong, P., Li, M., Okamura, A. M., and Hager, G. D. (2003). Spatial motion constraints: theory and demonstrations for robot guidance using virtual fixtures. volume 2, pages 1954–1959 vol.2. IEEE.

[Marayong and Okamura, 2004] Marayong, P. and Okamura, A. M. (2004). Speed-accuracy characteristics of human-machine cooperative manipulation using virtual fixtures with variable admittance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(3):518–532.

[McRuer and Jex, 1967] McRuer, D. and Jex, H. (1967). A review of Quasi-Linear pilot models. *Human Factors in Electronics, IEEE Transactions on*, HFE-8(3):231–249.

[Medina et al., 2012] Medina, J., Lee, D., and Hirche, S. (2012). Risk-sensitive optimal feedback control for haptic assistance. In *, 2012 IEEE International Conference on Robotics and Automation, 2012.*, pages 1025–1031.

[Meisel, 2011] Meisel, S. (2011). *Anticipatory Optimization for Dynamic Decision Making*, volume 51 of *Operations Research/Computer Science Interfaces Series*. Springer, 1 edition.

[Mihelj and Podobnik, 2012] Mihelj, M. and Podobnik, J. (2012). *Haptics for Virtual Reality and Teleoperation*, volume 64 of *Intelligent Systems, Control and Automation: Science and Engineering*. Springer Netherlands.

[Mobasser and Hashtrudi-Zaad, 2008] Mobasser, F. and Hashtrudi-Zaad, K. (2008). Transparent rate mode bilateral teleoperation control. *The International Journal of Robotics Research*, 27(1):57–72.

[Montague and Berns, 2002] Montague, P. and Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36(2):265–284.

[Mörtl et al., 2012] Mörtl, A., Lawitzky, M., Kucukyilmaz, A., Sezgin, M., Basdogan, C., and Hirche, S. (2012). The Role of Roles: Physical Cooperation between Humans and Robots. *The International Journal of Robotics Research*.

[Márquez, 2003] Márquez, H. (2003). *Nonlinear Control Systems: Analysis and Design*. Wiley.

[Naerum and Hannaford, 2009] Naerum, E. and Hannaford, B. (2009). Global transparency analysis of the lawrence teleoperator architecture. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 4344–4349.

[Nedic et al., 2008] Nedic, A., Tomlin, D., Holmes, P., Prentice, D., and Cohen, J. (2008). A simple decision task in a social context: Experiments, a model, and preliminary analyses of behavioral data. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 1115 –1120.

[Niemeyer et al., 2008] Niemeyer, G., Preusche, C., and Hirzinger, G. (2008). Telerobotics. In Siciliano, B. and Khatib, O., editors, *Springer Handbook of Robotics*, pages 741–757. Springer.

[Niemeyer and Slotine, 1991] Niemeyer, G. and Slotine, J.-J. (1991). Stable adaptive teleoperation. *Oceanic Engineering, IEEE Journal of*, 16(1):152–162.

[Nilsson and Riedel, 2004] Nilsson, J. W. and Riedel, S. (2004). *Electric Circuits*. Prentice Hall, 7 edition.

[Nitsch et al., 2012] Nitsch, V., Farber, B., Hellings, A., Jorg, S., Tobergte, A., and Konietschke, R. (2012). Bi-modal assistance functions and their effect on user perception and movement coordination with telesurgery systems. In *Haptic Audio Visual Environments and Games (HAVE), 2012 IEEE International Workshop on*, pages 32–37.

[Nudehi et al., 2005] Nudehi, S., Mukherjee, R., and Ghodoussi, M. (2005). A shared-control approach to haptic interface design for minimally invasive telesurgical training. *IEEE Transactions on Control Systems Technology*, 13(4):588–592.

[Oguz et al., 2010] Oguz, S., Kucukyilmaz, A., Sezgin, T., and Basdogan, C. (2010). Haptic negotiation and role exchange for collaboration in virtual environments. In *Haptics Symposium, 2010 IEEE*, pages 371 –378.

[Oguz et al., 2012] Oguz, S., Kucukyilmaz, A., Sezgin, T., and Basdogan, C. (2012). Supporting negotiation behavior with haptics-enabled human-computer interfaces. *IEEE Transactions on Haptics*, 5(3):274–284.

[Ortega et al., 1998] Ortega, R., Perez, J. A. L., Nicklasson, P. J., and Sira-Ramirez, H. (1998). *Passivity-based Control of Euler-Lagrange Systems: Mechanical, Electrical and Electromechanical Applications*. Springer, softcover reprint of the original 1st ed. 1998 edition edition.

[Passenberg and Buss, 2010] Passenberg, C. Peer, A. and Buss, M. (2010). A survey of environment-, operator-, and task-adapted controllers for teleoperation systems. *Mechatronics*, 20(7):787 – 801. Special Issue on Design and Control Methodologies in Telerobotics.

[Passenberg et al., 2013] Passenberg, C., Glaser, A., and Peer, A. (2013). Exploring the design space of haptic assistants: the assistance policy module. *IEEE Transactions on Haptics*, Early Access Online.

[Passenberg et al., 2011] Passenberg, C., Groten, R., Peer, A., and Buss, M. (2011). Towards real-time haptic assistance adaptation optimizing task performance and human effort. In *World Haptics Conference (WHC), 2011 IEEE*, pages 155–160.

[Patton and Klinkhieo, 2010] Patton, R. and Klinkhieo, S. (2010). LPV fault estimation and FTC of a two-link manipulator. In *American Control Conference (ACC), 2010*, pages 4647–4652.

[Peer and Buss, 2008a] Peer, A. and Buss, M. (2008a). A new admittance-type haptic interface for bimanual manipulations. *IEEE/ASME Transactions on Mechatronics*, 13(4):416–428.

[Peer and Buss, 2008b] Peer, A. and Buss, M. (2008b). Robust stability analysis of a bilateral teleoperation system using the parameter space approach. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 2350–2356. IEEE.

[Pezzementi et al., 2007] Pezzementi, Z., Okamura, A., and Hager, G. (2007). Dynamic guidance with pseudoadmittance virtual fixtures. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 1761 –1767.

[Powell and O'Malley, 2011] Powell, D. and O'Malley, M. (2011). Efficacy of shared-control guidance paradigms for robot-mediated training. In *World Haptics Conference (WHC), 2011 IEEE*, pages 427 –432.

[Powell and O'Malley, 2012] Powell, D. and O'Malley, M. (2012). The task-dependent efficacy of shared-control haptic guidance paradigms. *IEEE Transactions on Haptics*, 5(3):208–219.

[Raju et al., 1989] Raju, G., Verghese, G., and Sheridan, T. (1989). Design issues in 2-port network models of bilateral remote manipulation. In *, 1989 IEEE International Conference on Robotics and Automation, 1989. Proceedings*, pages 1316–1321 vol.3.

[Reed and Peshkin, 2008] Reed, K. B. and Peshkin, M. A. (2008). Physical collaboration of Human-Human and Human-Robot teams. *Haptics, IEEE Transactions on*, 1(2):108 –120.

[Rosenberg, 1993] Rosenberg, L. B. (1993). Virtual fixtures: Perceptual tools for telerobotic manipulation. In *Virtual Reality Annual International Symposium, 1993., 1993 IEEE*, page 76–82.

[Salcudean and Zhu, 2000] Salcudean, S. and Zhu, W.-H. (2000). Stability guaranteed teleoperation: an adaptive motion/force control approach. *IEEE Transactions on Automatic Control*, 45(11):1951–1969.

[Secchi et al., 2007] Secchi, C., Stramigioli, S., and Fantuzzi, C. (2007). *Control of interactive robotic interfaces a port-Hamiltonian approach*. Springer, Berlin; London.

[Shahdi and Sirouspour, 2009] Shahdi, A. and Sirouspour, S. (2009). Adaptive/Robust control for time-delay teleoperation. *IEEE Transactions on Robotics*, 25(1):196–205.

[Shen et al., 2004] Shen, J., Ibanez-Guzman, J., Ng, T. C., and Chew, B. S. (2004). A collaborative-shared control system with safe obstacle avoidance capability. In *2004 IEEE Conference on Robotics, Automation and Mechatronics*, volume 1, pages 119–123 vol.1.

[Sheridan, 1992] Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. MIT Press, Cambridge, MA, USA.

[Shorten et al., 2004] Shorten, R. N., Mason, O., O'Cairbre, F., and Curran, P. (2004). A unifying framework for the siso circle criterion and other quadratic stability criteria. *Int. Journal of Control*, 77(1):1–8.

[Sirouspour and Shahdi, 2006] Sirouspour, S. and Shahdi, A. (2006). Model predictive control for transparent teleoperation under communication time delay. *IEEE Transactions on Robotics*, 22(6):1131–1145.

[Spong et al., 2005] Spong, M. W., Hutchinson, S., and Vidyasagar, M. (2005). *Robot Modeling and Control*. Wiley.

[Stewart et al., 2010] Stewart, A., Cao, M., and Leonard, N. E. (2010). Steady-state distributions for human decisions in two-alternative choice tasks. In *American Control Conference (ACC), 2010*, page 2378–2383.

[Stewart et al., 2012] Stewart, A., Cao, M., Nedic, A., Tomlin, D., and Leonard, N. (2012). Towards Human&Robot teams: Model-based analysis of human decision making in two-alternative choice tasks with social feedback. *Proceedings of the IEEE*, 100(3):751–775.

[Strassberg et al., 1993] Strassberg, Y., Goldenberg, A. A., and Mills, J. K. (1993). Stability analysis of a bilateral teleoperating system. *Journal of Dynamic Systems, Measurement, and Control*, 115(3):419.

[Torres, 2007] Torres, B. (2007). *Control bilateral de robots teleoperados por convergencia de estados*. Tesis (Doctoral), E.T.S.I. Industriales (UPM).

[Tsuzuki and Guo, 2004] Tsuzuki, T. and Guo, F. Y. (2004). A stochastic comparison-grouping model of multialternative choice: Explaining decoy effects. In *Proceedings of the twenty-sixth annual conference of the cognitive science society*, pages 1351–1356. Citeseer.

[Ueberle et al., 2004] Ueberle, M., Mock, N., and Buss, M. (2004). VISHARD10, a novel hyper-redundant haptic interface. In *12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2004. HAPTICS '04. Proceedings*, pages 58–65.

[Unterhinninghofen, 2009] Unterhinninghofen, U. (2009). *Dynamic Assist Functions in Haptic Telepresence*. PhD thesis, Technischen Universität München.

[Veras et al., 2008] Veras, E. J., De Laurentis, K., and Dubey, R. (2008). Design and implementation of visual-haptic assistive control system for virtual rehabilitation exercise and teleoperation manipulation. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4290–4293.

[Vu and Morgansen, 2008] Vu, L. and Morgansen, K. (2008). Modeling and analysis of dynamic decision making in sequential two-choice tasks. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 1121 –1126.

[Weber et al., 2009] Weber, C., Nitsch, V., Unterhinninghofen, U., Farber, B., and Buss, M. (2009). Position and force augmentation in a telepresence system and their effects on perceived realism. In *EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*, pages 226–231.

[Woodruff et al., 2012] Woodruff, C., Vu, L., Morgansen, K., and Tomlin, D. (2012). Deterministic modeling and evaluation of decision-making dynamics in sequential two-alternative forced choice tasks. *Proceedings of the IEEE*, 100(3):734 –750.

[Yokokohji and Yoshikawa, 1994] Yokokohji, Y. and Yoshikawa, T. (1994). Bilateral control of master-slave manipulators for ideal kinesthetic coupling-formulation and experiment. *Robotics and Automation, IEEE Transactions on*, 10(5):605 –620.

[Yu et al., 2003] Yu, H., Spenko, M., and Dubowsky, S. (2003). An adaptive shared control system for an intelligent mobility aid for the elderly. *Autonomous Robots*, 15(1):53–66.

[Zhu and Salcudean, 2000] Zhu, W.-H. and Salcudean, S. (2000). Stability guaranteed teleoperation: an adaptive motion/force control approach. *IEEE Transactions on Automatic Control*, 45(11):1951–1969.