

# Progress in Biomedical Engineering



## TOPICAL REVIEW

# A review on interaction control for contact robots through intent detection

### OPEN ACCESS

RECEIVED  
7 February 2022

REVISED  
1 June 2022

ACCEPTED FOR PUBLICATION  
15 July 2022

PUBLISHED  
9 August 2022

Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Yanan Li<sup>1,6</sup>, Aran Sena<sup>2,6</sup> , Ziwei Wang<sup>2</sup>, Xueyan Xing<sup>1</sup>, Jan Babič<sup>3,4</sup>, Edwin van Asseldonk<sup>5</sup> and Etienne Burdet<sup>2,\*</sup>

<sup>1</sup> Department of Engineering and Design, University of Sussex, Brighton BN1 9RH, United Kingdom

<sup>2</sup> Department of Bioengineering, Imperial College of Science, Technology and Medicine, London W12 0BZ, United Kingdom

<sup>3</sup> Laboratory for Neuromechanics and Biorobotics, Department of Automation, Biocybernetics and Robotics, Jožef Stefan Institute, Ljubljana, Slovenia

<sup>4</sup> Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

<sup>5</sup> Faculty of Engineering Technology, University of Twente, 7500 AE Enschede, The Netherlands

<sup>6</sup> Equal contribution.

\* Author to whom any correspondence should be addressed.

E-mail: [e.burdet@imperial.ac.uk](mailto:e.burdet@imperial.ac.uk)

**Keywords:** control, estimations, adaptations, interactive, robot, human, learning

## Abstract

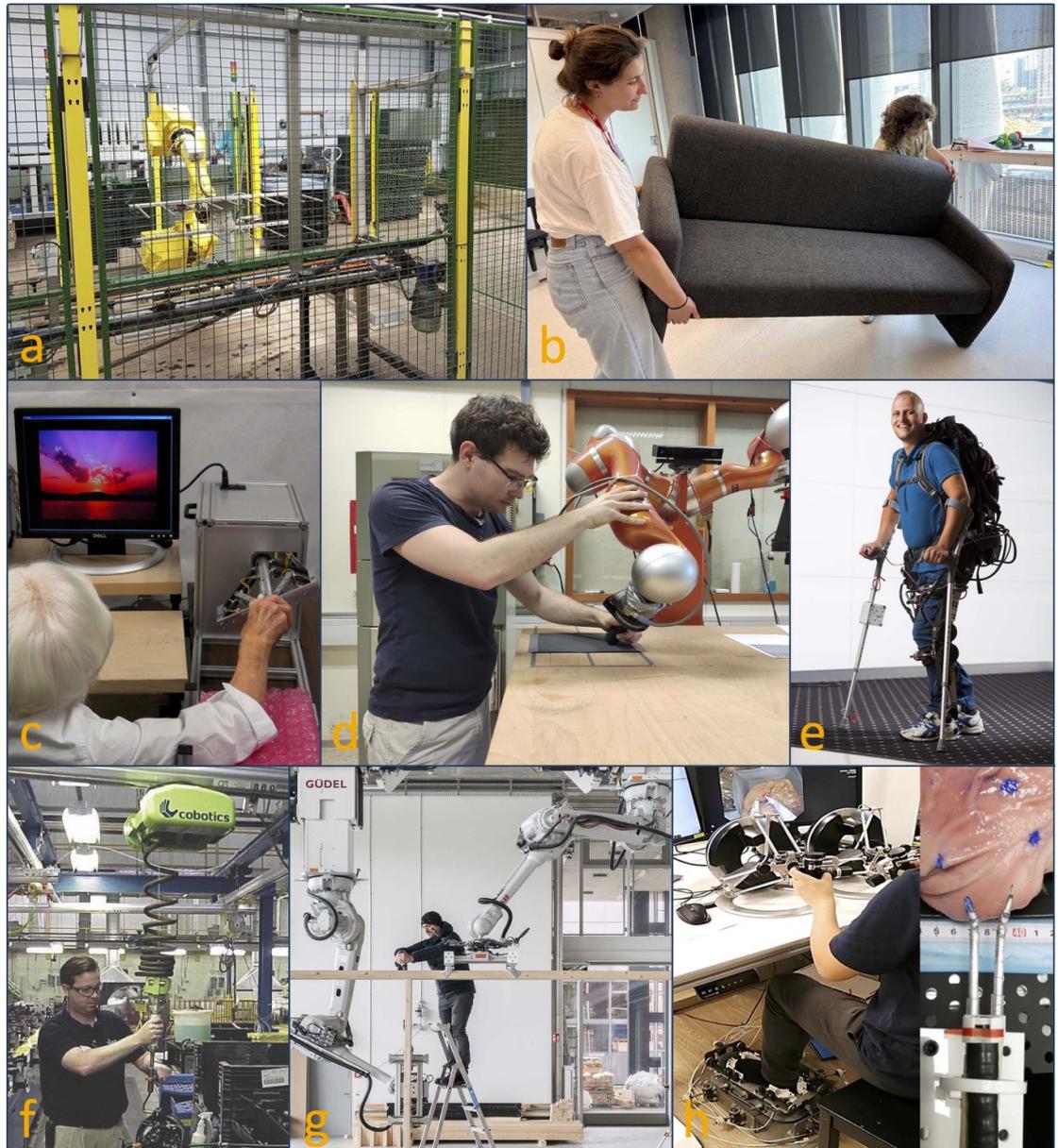
Interaction control presents opportunities for *contact robots* physically interacting with their human user, such as assistance targeted to each human user, communication of goals to enable effective teamwork, and task-directed motion resistance in physical training and rehabilitation contexts. Here we review the burgeoning field of interaction control in the control theory and machine learning communities, by analysing the exchange of haptic information between the robot and its human user, and how they share the task effort. We first review the estimation and learning methods *to predict the human user intent* with the large uncertainty, variability and noise and limited observation of human motion. Based on this motion intent core, typical *interaction control strategies* are described using a homotopy of shared control parameters. Recent methods of haptic communication and game theory are then presented *to consider the co-adaptation of human and robot control* and yield versatile interactive control as observed between humans. Finally, the limitations of the presented state of the art are discussed and directions for future research are outlined.

## 1. Introduction

Traditional robotic manipulators, such as those used in the automotive industry, are potentially dangerous to human workers and thus are kept separated from them (figure 1(a)). Inspired by the benefits of human–human interaction ([1], figure 1(b)), robotic systems have been increasingly used to work in physical interaction with humans (figures 1(c)–(h)). In these *contact robots*, the physical interaction arises directly, or through an object such as when carrying a large object together. Contact robots offer a broad range of new opportunities for application in various fields such as manufacturing, microfabrication, or shared driving [2], as well as in a number of medical contexts, such as surgery [3] and neurorehabilitation [4].

For instance, teleoperation can enable an operator to carry out actions in remote, inaccessible or hazardous environments, such as long-distance space exploration activities, micro/nano-world manipulation or surgical operations. Robot systems designed for collaborations with human workers, often termed collaborative robots, can carry out tasks that are difficult to perform ergonomically when unassisted or otherwise uneconomical to automate [5, 6], or that require close proximity to human operators [7]. Robotic systems are also developed to facilitate physical rehabilitation [8–12] and sport training.

Historically, human–robot interaction consisted of developing safe strategies to use a robot in the proximity of humans, but not in contact with them [13]. For instance, the presence and location of humans detected from machine vision or infrared maps can be used to determine safe areas for the robot to move in

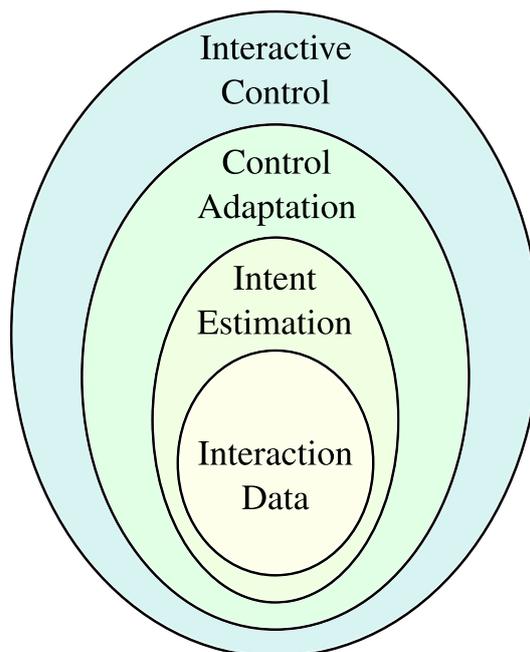


**Figure 1.** Examples of interaction control. (a) Traditional industrial manipulators are separated from human workers. (b) Humans can carry out tasks efficiently together by relying on haptic communication. (c) Robot-assisted neurorehabilitation of the hand function (from National University of Singapore and Imperial College London). (d) Collaborative robot for manufacturing (from Jožef Stefan Institute, Ljubljana, Slovenia). (e) Using the Symbion exoskeleton (from University of Twente) to facilitate mobility. (f) Intelligent gravity support for ergonomic manufacturing (reproduced with permission from EJ Colgate). Reproduced with permission from EJ Colgate. (g) Human–robot interaction for versatile operation in construction, showing two robots assisting in bolting the timber beams. Reproduced with permission. A project member manually bolts the timber beams together that have been jointly preplaced by the two robots. Photo: NCCR Digital Fabrication/Roman Keller. (h) Teleoperated ‘three hand surgery’ using a soft endoscope equipped with tools (from Nanyang Technical University, Singapore, reproduced with permission).

and avoid collisions. To better decide how to safely and efficiently assist a human operator the robot can implement predictive assistive behaviours, rather than simply engaging with reactive behaviours, for example by inferring the operator’s motion target intent to help guide their movement to the goal.

This intent can be inferred through observation of the operator’s motion during the interaction, and may also incorporate prior knowledge of the task being performed. In tasks requiring physical contact between the human and robot, sensory cues such as interaction forces, motion and electrophysiological activities can provide information about the human operator’s intents, that are unobservable for a purely vision-guided interaction.

Furthermore, recent studies have revealed a number of performance benefits when humans physically interact and collaborate while performing a common task [1, 14]. In particular, it has been shown that human partners share their motion plan through haptic sensing (the synthesis of touch and force sensing),



**Figure 2.** Control for contact rich human–robot interactions can be described in terms of layered dependencies, with each layer building on the last to enhance the overall interactive capability of the robot system. At the core is the sensor data of the interaction and prior information. This data can then be processed into intents and task inferences, that can then be used for control adaptations and interactive behaviours.

which they use to improve own performance [15–17]. This highlights the potential for contact robots to rely on *haptic communication*, the exchange of haptic information, to identify the human operator motion plan and improve interaction control [18, 19]. Dedicated techniques developed in signal processing and machine learning (ML), such as linear quadratic estimation (LQE), dynamic movement primitives (DMP), Gaussian processes, etc, can be used to identify a partner’s control.

Another major aspect of interaction control concerns *how to share the effort* between the robot and its user. Many interaction control methods tend to assign fixed roles of leader and follower to the human and the robot (such as in teleoperation [20]) or conversely to provide a fixed dynamic environment imposed on the user (such as in robot-assisted physical training [21]). However, while the human will automatically learn the robot control scheme and adapt to it [15, 22], the robot may also adapt its control to the human based on knowledge acquired through haptic communication. By understanding the human user control and using appropriate control theory or ML algorithms, the robot may develop appropriate interaction strategies to share effort considering their respective strengths [23, 24].

This paper reviews techniques for the interaction control of a contact robot with its human operator, focusing on the mechanical interaction and the exchange of haptic information (through the interaction force and tactile information). Compared to previous recent review papers (such as [23, 25]), this paper uses a systematic framework based on a dynamic model, cost functions, and control laws to compare various algorithms and describe future control designs. It also considers human neuromechanical control [26] as a key factor to intent estimation when interacting with a human partner. This includes recent studies on the physical interaction between humans that reveal how haptic communication is used to model the partner and improve interaction control [1, 15, 27]. These results underline the critical role of motion intent to shape the interaction between two agents.

The techniques of interaction control for contact robots can be described as an onion structure, with layers built around the core of its human user’s intention detection (figure 2). Based on interaction data and a-priori information, a contact robot may read the motion/control intent. This information can be used to adapt the robot’s behaviour unilaterally, or to develop interactive control considering the human natural co-adaptation.

The sections in the rest of the paper correspond to this structure. Section 2 formalizes the interaction control dynamics framework between a robot and its human user. The physical model of the human–robot system interacting with the environment forms the basis for how state updates through sensing are used to adapt the robot behaviour. The systems and techniques to identify the intent of the human operator are then described in section 3. Section 4 describes the various schemes developed during the last decades to adapt the

interaction in a contact robot. Section 5 then discusses tools to analyse human–robot co-adaptation and to create interactive control between these two agents. Finally in section 6 limitations of the current state of the art are analysed and future research directions are identified,

## 2. Human–robot control dynamics

### 2.1. Interaction dynamics

We consider physical human–robot interaction through a direct kinaesthetic interface (e.g. an exoskeleton), through a commonly held object (e.g. in co-manipulation), or through a haptic device (e.g. for teleoperation). All these interactions can be described using the dynamic model:

$$\mathbf{u} + \mathbf{u}_h = \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}), \quad (1)$$

expressing that the robot control command  $\mathbf{u}$  and human motor command  $\mathbf{u}_h$  drive the task dynamics  $\mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}})$ , where  $\mathbf{x}$  is position vector and all variables are functions of time. Depending on the context of the aforementioned different interactions,  $\mathbf{x}$  may represent the position of the collaboratively held object in co-manipulation, the position of the follower robot in teleoperation, or the joint angles of an exoskeleton. Moreover, although the above model describes interaction between a human and a robot, it can be extended to modelling of human–human or robot–robot interaction.

Without loss of generality, we assume that the robot control includes *feedforward* and *feedback* components:

$$\mathbf{u}(t) = \mathbf{v}(t) + \mathbf{w}(t), \quad (2)$$

where  $t$  represents the time. The feedforward component  $\mathbf{v}$  uses prior knowledge of the system acquired through learning or identification to compensate for the (typically nonlinear) task dynamics. The feedback component  $\mathbf{w}$  uses online evaluation of control performance to react in an appropriate way to unexpected situations:

$$\mathbf{w}(t) = \mathbf{K}\mathbf{e}(t) + \mathbf{B}\dot{\mathbf{e}}(t), \quad \mathbf{e}(t) = \mathbf{x}^*(t) - \mathbf{x}(t), \quad (3)$$

where  $\mathbf{x}^*$  is the motion target (a desired position or trajectory) and  $\mathbf{K}, \mathbf{B}$  are feedback gain matrices corresponding to stiffness and viscosity. Note that these feedback matrices can be updated to deal with uncertainties in the system, e.g. using adaptive control [28]. This gives a general form of feedforward and feedback components, which can be modified in different conditions and for different systems.

Human motor control can be modelled similarly:

$$\begin{aligned} \mathbf{u}_h &= \mathbf{v}_h(t) + \mathbf{w}_h(t), \\ \mathbf{w}_h(t) &= \mathbf{K}_h(\mathbf{u}_h) \mathbf{e}_h(t) + \mathbf{B}_h(\mathbf{u}_h) \dot{\mathbf{e}}_h(t), \\ \mathbf{e}_h(t) &= \mathbf{x}_h^*(t) - \mathbf{x}(t). \end{aligned} \quad (4)$$

In this case  $\mathbf{w}_h$  corresponds to the viscoelastic response to perturbations relative to the planned trajectory due to muscle mechanical properties and reflexes, where both the stiffness  $\mathbf{K}_h(\mathbf{u}_h)$  and viscosity  $\mathbf{B}_h(\mathbf{u}_h)$  increase with the motor command magnitude  $|\mathbf{u}_h|$  [26].

### 2.2. Estimating the human user's intent

If the human parameters are not considered by the robot and the human does not consider the robot, they are in a relationship of *co-activity* where each of the two agents considers only themselves [29]. Interaction via co-activity may work well if the robot and human carry out independent and complementary subtasks, e.g. an intelligent gravity support device enabling a human operator to move a workpiece in a horizontal plane with little effort (figure 1(f), [30]). However, when the robot and human work on the same task in an uncoordinated way, this may lead to inefficient interaction, instability or failure [31]. Therefore, it is critical for contact robots to consider the interaction dynamics with their users. To collaborate with a human operator, the robot should know or predict some elements of their movement or their control, usually referred to as *intent estimation* in physical human–robot interaction literature.

Estimating the human motion intent requires modelling their neuromechanical control. Let us consider a nonlinear stochastic framework recently introduced to model human motor control with feedforward and feedback components and considering the inherent sensorimotor noise [32, 33]. In this framework,

the feedforward control component  $\mathbf{v}$  corresponds to minimization of the expected cost (for visibility we drop the  $_h$  subscript in  $\mathbf{e}, \mathbf{v}$ ):

$$C_v = \mathbb{E} \left[ \int_0^T [\mathbf{e}(t)' \mathbf{Q} \mathbf{e}(t) + \mathbf{v}(t)' \mathbf{R} \mathbf{v}(t)] dt \right], \quad (5)$$

where  $\mathbf{Q}, \mathbf{R}$  are weighting matrices and  $T$  is the terminal time of the planned motion. This cost function expresses human sensorimotor control as a trade-off between information (through  $\mathbf{e}$ ) and energy (from  $\mathbf{v}$ ) [34].

The feedback component  $\mathbf{w}$  can be computed using a linearization of the human motor command dynamics, by minimizing the cost:

$$C_w = \int_0^{T_+} [\hat{\mathbf{e}}' \mathbf{Q} \hat{\mathbf{e}} + \mathbf{w}(t)' \mathbf{R} \mathbf{w}(t)] dt, \quad \hat{\mathbf{e}} \equiv \mathbf{x}_h^*(t) - \hat{\mathbf{x}}(t), \quad (6)$$

where  $T_+ > T$  indicates that task completion takes longer than the planned motion. For linear feedback, this optimal control method is called *linear quadratic regulation* (LQR). When the time window  $[0, T_+]$  moves forward, minimizing the cost  $C_w$  yields model predictive control, which can be used for nonlinear systems and include various constraints [35, 36]. In equation (6) the current movement  $\hat{\mathbf{x}}(t)$  is estimated by minimizing the square error prediction:

$$C_x = \mathbb{E} [ \|\mathbf{x}(t) - \hat{\mathbf{x}}(t)\|^2 ]. \quad (7)$$

For a linear system this yields *linear quadratic estimation* (LQE) or *Kalman filtering*, which, combined with LQR (assuming independent Gaussian noise), yields *linear quadratic Gaussian control*.

This model can predict human arm movements in various visual and dynamic conditions [32, 33]. However it requires knowledge of the neuromechanical parameters that will depend on the user and on their action. The many parameters involved in such complex model can only be indirectly observed by the robot through interaction responses or biosignal analysis, and would be difficult to identify during movement. On the other hand, it is possible to extract information from the human user in realtime directly from interaction data without a formal model. For instance, the data-driven technique of [37] could accurately predict a simple arm reaching movement for over 300 ms in advance, by comparing the ongoing movement with individual movement patterns. However it does not generalize to other movements.

Therefore, various approaches between these two model-based and data-driven extremes have been developed to learn the human user's motion intent. These methods, based on signal processing, ML and control theory, will be presented in section 3. They represent attempts to address the trade-off between a prediction based on simple models with few parameters that can be efficiently identified and run in realtime (but have limited generalization ability), and on more complex models offering the versatility necessary to be used for various tasks (but are difficult to realize).

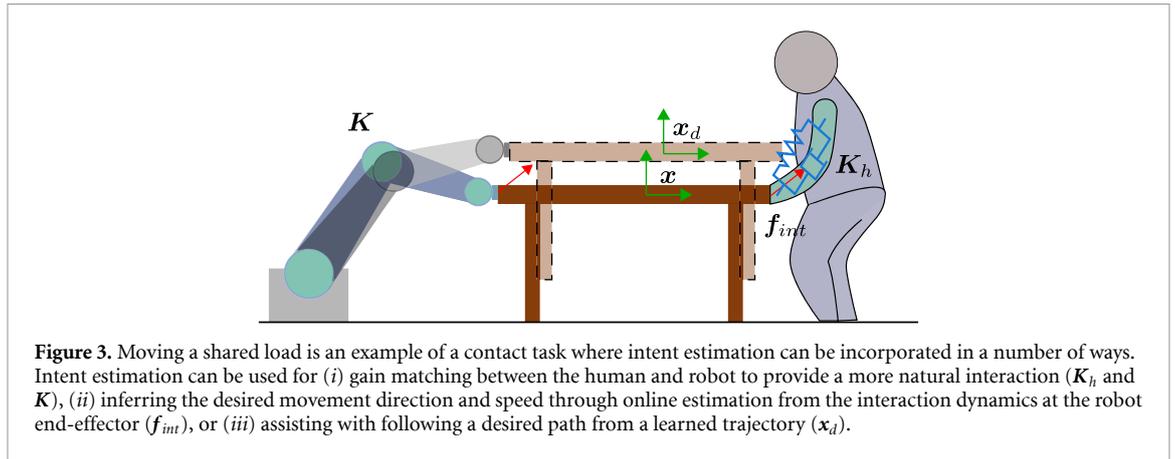
### 2.3. Human–robot interaction control

How should the robot consider detected information on the human user's motion and share the control with them? For instance, in the case of exoskeleton control, it can be desirable to yield 'transparent' interaction by offering little resistance, or to provide assistance along the user's planned movement. In contrast, to avoid reaching vital areas, a surgical robot may have to take the lead in order to keep the tool away from these areas [38]. Different tasks will require interaction behaviours between these extremes, defined according to a task-specific higher-order logic [39], which we describe by:

- desired task objectives, e.g. stable interaction, reaching a task target  $\mathbf{x}_d$  (that can be different from  $\mathbf{x}^*$ ) or tracking a desired force trajectory  $\mathbf{f}_d$ ;
- a desired (fixed or varying) relationship between  $\mathbf{u}, \mathbf{K}, \mathbf{B}, \mathbf{x}^*$  and  $\mathbf{u}_h, \mathbf{K}_h, \mathbf{B}_h, \mathbf{x}_h^*$ .

Using these control notations of section 2.1, section 4 will describe typical approaches for contact robots to adapt to the human user developed in the literature.

The robot control should further consider the adaptation of the human to dynamic environments [40, 41] or the interaction to human partners as was observed in [1, 16]. The bilateral adaptation between human and robot (or other pair of agents such as human–human and robot–robot) can be modelled/designed using *game theory* (GT) as reported in section 5.



### 3. Human intent estimation

Intent estimation can refer to many methods and approaches, encompassing topics such as activity recognition, safety systems, etc. In physical human–robot interactions, using intent detection to generate a forward prediction of an ongoing or future motion enables the robot to *actively* participate in the interaction. This enables many new behaviours over passive interactions, such as compensating for inertial effects to reduce the effort required from the user to move with the robot, for instance when using an exoskeleton system [42]. We consider here three categories of intent estimation that are particularly relevant to the modelling and control of physical human–robot interactions:

- motion intent* from movement kinematics and system dynamics, e.g. inferring the intended motion target so that the robot can actively guide the human towards the goal;
- action intent* from discrete action sets, e.g. for selecting an appropriate grasp type for a prosthetic based on electromyography (EMG) signals;
- task-level intent*, e.g. given knowledge of the overall goal task, a robot can provide corrective actions that contradict the human’s current intended motion, as might be used in a rehabilitation context.

A mixture of these approaches may be used within one system to manage different aspects of a task; however, these categories partition the different approaches reasonably well based on the underlying methods used, which range from control and signal processing methods to various ML methods.

#### 3.1. Motion intent estimation

Motion intent estimation must involve some form of assumption over the human movement and/or use some state estimation technique to provide a forward prediction of the current motion, typically over a relatively short time horizon, with the goal of directly adapting the robot control dynamics to better suit the interaction.

Early work on motion and dynamics based intent estimation considered collaborative transport problems, where a human–robot team is tasked with moving an object together ([43, 44], figure 3). Earlier modelling of this shared movement assumed human motion during the interaction to be maximally smooth, as in the average of reaching arm movements, and can be modelled as a fifth order polynomial (minimizing the magnitude of time derivative of the acceleration) [45]. More recent works have sought to apply this model for human–robot interaction, by considering individual motion characteristics [46]. While this model presents a convenient parameterization, its validity is limited to reaching arm movements, and it fails at accurately predicting movements with constraints, such as when picking/depositing an object, or during interaction with the environment [26].

Other approaches to motion intent estimation for contact robots are based on the interaction force/torque  $u_h$  which can be measured through a sensor at the interface [47]. Using the model structure of section 2.1, the human control input can be defined by (i) virtual control gains and (ii) a virtual human target  $x_h^*$  representing the motion intent:

$$w_h = -K_h(x - x_h^*) - B_h\dot{x}. \quad (8)$$

In this expression, the virtual gains  $K_h, B_h$  can be measured or computed from human neuromechanical models [26], learned using average population data, or just assumed a practical value, e.g. to match the robot control gains.  $x_h^*$  can then be estimated from this equation using LQE.

Data driven methods have been considered to estimate the human reaching target during a physical human–robot interaction limb guiding task. In [48], a radial basis function neural network was inserted in the impedance control loop to compensate for the unknown dynamics of the human limb in contact with the robot in an online-adaptive manner, which provides an estimate of the human motion target. ML techniques for biosignal analysis can be used for human motion intent prediction. For instance, EMG signals can be used to estimate human motion in 3D space, with recurrent neural network architectures yielding better results than a support vector regression or a non-recurrent network architecture [49]. In [50], a deep convolutional neural network was used to estimate interaction force from EMG, potentially avoiding the need to incorporate interaction force sensors to facilitate human–robot physical interaction. However, prediction based on EMG signals will suffer from their large inter-trials variability and noise [26].

### 3.2. Action intent estimation

Action intent estimation can be used to modulate the type of human–robot interaction, by providing a discrete action set that the robot can choose from based on observations of the human user. It aims to enable more complex and task-specific interactions that would be impossible by only considering the interaction dynamics as found in motion intent estimation. This form of intent is typically concerned with a longer time horizon than motion intent estimation, and generally requires a higher logic beyond (not represented in the notation of section 2.1).

For example, a subacute stroke patient may not be able to open the hand alone and thus want to control a hand exoskeleton system with their forearm EMG signals. Their EMG signals may be degraded so could not be used to derive a continuous control command to the robot, but a simple thresholding method is generally sufficient to infer the intended action [51]. In comparison, EEG has also been used in this study, but did not exhibit superior performance and is more complex to set up. Another example is when the robot body is covered with an electronic tactile sensing skin which the human user can touch to elicit ‘event-like’ actions [52, 53].

Other recent approaches have applied ML methods to classify intended actions. This might involve using sensing which is not directly relevant to the main task, but offers a communication route between the human and robot. An example of this is processing bicep and tricep EMG signals with a multilayer perceptron to yield discrete ‘move up’ and ‘move down’ control signals during a collaborative lifting task (figure 3) [54]. This allows the robot to interact with the human even while lifting a completely compliant object, e.g. a loose piece of rope, where there would not be a dynamic interaction at the human–robot interface for motion intent estimation approaches, without applying a stretching force on the rope.

### 3.3. Task intent estimation

Intent estimation at the task-level can give a robot partner prior knowledge of the overall task or goal that they should try to achieve, enabling further complex human–robot interactions. Rather than considering the next moment or action as seen in motion or action intent estimation, task-level intent estimation may consider a full expected trajectory representing an objective behaviour the robot should try replicated with their partner.

Consider a human–robot team performing a peg-in-hole type task. One approach to modelling the control requirements for completing this task is to assume that accurate positioning of the robot end-effector is required specifically at the insertion stage of the task, but not necessarily while initially moving the robot to the approximate region of the hole. If the robot were to have a high control gain through out the task, it would be possible for the human to position the robot accurately, but not do this particularly quickly or with low-effort, as the high-level of damping used in the robot controller would resist changes in motion. Some control heuristics can mitigate the issue, for example a robot could provide high gain during slow movements to help with accuracy [55]; however, this will limit movement in *all* directions and will not necessarily guide the user. Instead, task intent estimation can allow for varying controller gains in a *task-relevant* manner [56–61]. In the case of [58], for a peg-in-hole type task a minimal intervention controller is defined based on a learned task representation that results in the robot behaving compliantly during free-space movements while moving toward the hole, then engaging in a task-directed stiffness as it performs the insertion stage of the task, offering resistance in directions of motion that do not align the peg co-axially with the hole, with the overall effect of reducing both robot and human effort.

Providing the robot with knowledge of the task-level intent also enables the robot to act autonomously, creating opportunities for *robot adaptation*, for example as a means of compensating for noisy communications during teleoperation by providing a fall-back behaviour [62], or as a means of gradually

taking over execution of the task as the robot system gains confidence in the learned task representation [63–65]. Alternating between human control and robot autonomy is discussed further in section 4.

A representation of a task can simply be a known and fixed trajectory, but it is often required from a robot to *learn* a representation of the task that can be adapted to different conditions. There is a large body of prior work detailing how such representations can be learned to enable robots to perform new tasks [57, 66–70]. A general overview is presented here, along with a description of how such representations can be used for intent estimation in physical human–robot interactions. As summarised in [70], learning a task specifically involves learning a policy, which is a mapping:

$$\pi : (\mathbf{s}, \mathbf{x}_s) \rightarrow \mathbf{a}, \quad (9)$$

between observed states  $\mathbf{x}_s$  and/or contexts  $\mathbf{s}$  to an action  $\mathbf{a}$ . In the context of robot task learning for intent estimation, these output actions can represent different levels of action, from (a) low-level actions, such as direct motor commands, motion targets, or system gains, to (b) full trajectories of these low-level actions,  $\boldsymbol{\tau}$ , or (c) abstract high-level task actions that represent a series of trajectories to perform some sequence of tasks, e.g. ‘pick up the cup’. A policy can either be deterministic,  $\mathbf{a} = \pi(\mathbf{s}, \mathbf{x}_s)$ , or stochastic, where actions are drawn from a policy distribution conditioned on current states and contexts,  $\mathbf{a} \sim \pi[\mathbf{a} | (\mathbf{s}, \mathbf{x}_s)]$ .

There are many approaches for deriving a policy for robot control. ML for robot control most commonly considers *behavioural cloning* and *reinforcement learning* (RL).

### 3.3.1. Behavioural cloning

Describes methods where the robot uses data samples of input observed states and the corresponding output actions to directly produce a map between states and actions as described in equation (9).

In systems where labelled data sets for a task are provided through physical interaction with a human partner, a behavioural cloning system might also be described as a *learning from demonstration* system [67]. Here, a dataset of  $N$  demonstrations,  $\mathcal{D}$ , is provided. The demonstration dataset consists of  $N$ -trajectories,  $\boldsymbol{\tau}$ , contexts,  $\mathbf{s}$ , and optionally reward signals,  $\mathbf{r}$ , that specify the performance of the observed trajectory in the context of the goal task,  $\mathcal{D} = \{(\boldsymbol{\tau}_i, \mathbf{s}_i, \mathbf{r}_i)\}_{i=1}^N$ . Each trajectory is defined by a sequence of features,  $\boldsymbol{\tau} = [\phi_0, \dots, \phi_T]$ , representing the system state at each time step. A common assumption for learning a policy from human-provided demonstrations is that the observed human-provided features represent the true goal policy, i.e. an *expert* policy. The robot learner’s goal is then to minimize the difference between the distribution over the observed expert features,  $\mathbf{q}(\phi)$ , and the distribution of features observed from the learner’s policy,  $\mathbf{p}(\phi)$ :

$$\pi^* = \operatorname{argmin} D[\mathbf{q}(\phi), \mathbf{p}(\phi)], \quad (10)$$

where  $D$  is some similarity measure between the expert and learner feature distributions, where the learner’s policy is typically learned using some supervised learning method [70]. An example that follows this paradigm and is commonly used in robot learning systems is the *Gaussian Mixture Model*. The GMM is a general probabilistic learning method which encodes the provided data set as a joint distribution between the input and output states,  $p(x, y)$ . Actions can be generated from this joint distribution through a Gaussian Mixture Regression which computes the conditional distribution for the output given an input vector  $p(y|x)$  [57, 58]. While behavioural cloning generally does not involve the learning agent generating new data for learning in undemonstrated regions (as found in RL), many methods have been developed to represent underlying invariances in a task with the learned model such that the robot can generate effective trajectories in previously unobserved task conditions [58, 71, 72]. An additional advantage to a GMM-based representation is that the encoded covariance data may be used for modulating the robot impedance in a task-directed manner [58]. These methods have been applied to a variety of tasks with contact robots, including collaborative object movement [73], converting human-human demonstrations to human-robot task policies [74].

A common alternative approach to probabilistic methods for robot task learning uses *motion primitives* to represent tasks. In particular, *dynamic movement primitives* (DMP) yields a method that uses a single demonstration to encode a task as a dynamical system such that the robot will converge to a specified goal state from the current state [75]. The task-specific trajectory is incorporated to the dynamical system through a learned forcing function, which seeks to reproduce the provided demonstration exactly. Examples of DMPs in the field of contact robots include object-handover tasks [76], and predictive modelling of human motion for robust exoskeleton control [77]. Recent works have considered extensions and combinations of probabilistic, motion primitive, and kernel based learning methods, to address issues such as high-dimensional state spaces, e.g. with *Kernelised Movement Primitives* [78, 79].

### 3.3.2. Reinforcement learning (RL)

While there has been a number of developments to improving the generalization of behaviour cloning, these supervised learning methods are typically not robust in conditions that deviate from the demonstrated conditions [70]. RL presents an alternate paradigm to task learning that enables a robot to learn a policy for a sequential decision making process where the environment dynamics are unknown [80, 81]. Generally, RL involves the robot interacting with the environment using its current policy, collecting state observations and possibly rewards, until it reaches some termination condition or state.

A key distinction from behavioural cloning methods is that this process can be achieved without any data being provided to the robot on the task that must be performed. Instead the robot can perform explorative actions in the task space to discover an optimal policy. While there have been great advances in perception and learning state representations in the field of deep learning that have enhanced RL [82], the expense and risk involved in pure exploration based data collection for real-world human–robot interactions is a significant bottleneck for deploying RL in real-world robotics [68, 83].

Rather than beginning learning from scratch, task demonstrations can be incorporated to accelerate learning by reducing the policy search space through *offline* learning, often called *inverse reinforcement learning* (IRL) [84]. Rather than learning a direct mapping between states and actions, the learning process attempts to infer the demonstrator's reward function, such that the robot can learn a policy that can generalise more effectively through exploration. In addition to the reduction in computational cost, this can serve to reduce the risk involved with a pure unstructured exploration approach to determining a task policy for a robot that must come in to contact with the environment or a human partner [68].

As highlighted in [85], the availability of task-relevant data is critical for enabling effective RL to take place. As discussed above, pure exploration of a task policy in the real world is typically not feasible, given the high number of samples required and safety concerns around random policy exploration. Key areas of research that aim to address this issue even more efficiently than IRL are sim-to-real transfer learning, and zero-/one-/few-shot learning. In sim-to-real transfer learning, the goal is to learn a policy in simulation, where data is cheap and exploration is safe, then transfer this policy to a real robot system for deployment [86, 87]. Practically, this encounters many challenges due to the gap between simulated and real-world task environments. Few-shot learning then aims to address the challenges of deploying robots on new tasks in the real-world by leveraging policies learned for previously learned similar tasks [88–90].

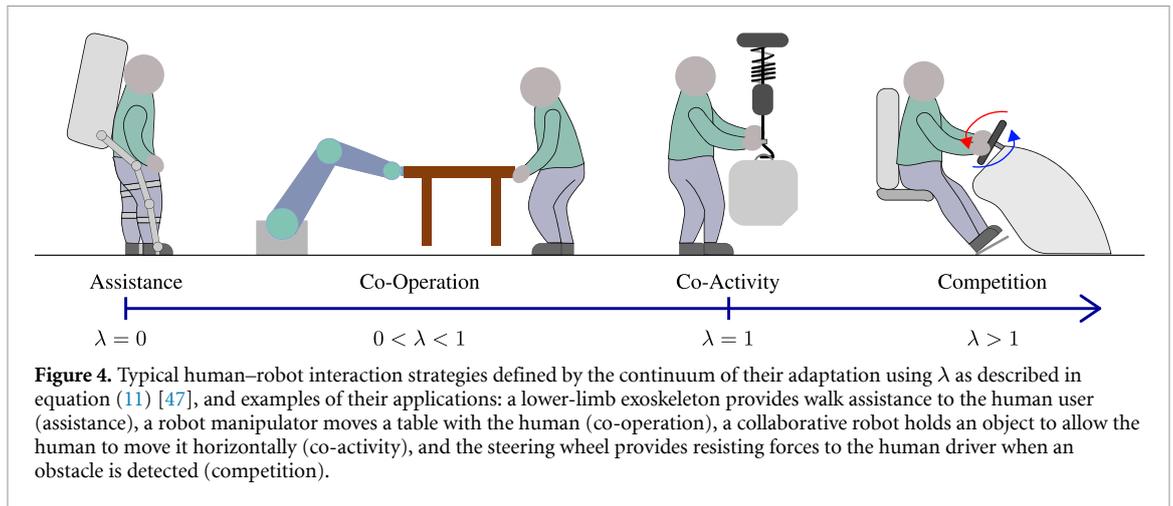
### 3.4. Mutual human–robot learning

With both behavioural cloning and RL based approaches to task learning for contact robots, the human operator should usually be the reference and thus have a leading role in the learning process. This can be done by using an incremental approach to modify a repeated motion with a GUI or through direct physical interaction [91–93], which can yield more accurate and smoother action. Many task learning methods can use the physical interaction with the human operator to transfer and refine expert demonstrations by directly pushing or pulling the contact robot along the desired task trajectories during behavioural cloning [67].

Physical interactions can also enable the human to provide online corrective actions during a RL process [94]. It is also possible to directly incorporate the coadaptation that takes place when a human and robot are learning to perform a new task collaboratively [95]. Here, a human and a robot collaborate to complete a ball maze task, where the teleoperated commands sent to the robot by the human are included in the RL state observations. Incorporating the human in the learning offers the opportunity for co-learning to occur, where as the human improves at the task the robot can adapt its own policy to better interact with them, and vice-versa.

When learning a control policy that involves interacting with a human partner that is also learning a policy, neglecting to consider the human adaptation could result in a *covariate shift* in the task distribution that makes the learned model ineffective at providing an accurate estimation of the human's intent [70, 95], unless this co-adaptation is explicitly accounted for. There is much scope for investigating the dynamic interaction that takes place between partners learning a task, with previous works modelling this problem between a human and a robot as a noisy communication process [96–98]. A game-theoretic control perspective is presented on this topic in section 5.

In addition to the challenge of co-adaptive learning, when considering learning processes that must directly interact with, and involve human partners, there are some key points to note. First, as data is collected through real-world interaction with a human partner, data collection is an expensive process. This results in the need for learning methods that are effective with limited data. Second, humans exhibit a wide range of physical properties that can make determining appropriate control commands for physical interactions difficult, particularly given the data collection limitations. It has been shown that population-level models of human motion can perform poorly, thus requiring model personalization [37, 99].



## 4. Robot control adaptation

Using elements of the human user’s control that can be identified as discussed in the previous section, the robot can adapt its control according to an interaction strategy. Using the modelling of section 2.1, this control adaptation can be implemented on the robot motion target, control input or on the control gains (figure 3), as will be presented in this section. Extending the approach of [100], we define a homotopy of human–robot interaction behaviours:

$$\xi = \xi_0 + (1 - \lambda)\xi_h, \quad \lambda \geq 0, \quad (11)$$

where  $(\xi, \xi_h)$  denote the robot and human targets ( $\mathbf{x}^*, \mathbf{x}_h^*$ ), control inputs ( $\mathbf{u}, \mathbf{u}_h$ ) or gains ( $[\mathbf{K}, \mathbf{B}], [\mathbf{K}_h, \mathbf{B}_h]$ ), and  $\xi_0$  is a nominal parameter. As illustrated in figure 4, this continuous parameter’s representation can describe the typical interaction behaviours introduced in [101] by varying  $\lambda$ .  $\lambda = 0$  corresponds to *Assistance*,  $\lambda = 1$  to *Co-activity*,  $0 < \lambda < 1$  to *Cooperation* and  $\lambda > 1$  to *Competition*. For example, when  $\xi$  stands for motion target in a load sharing task in figure 3, competition refers to the case where human’s and robot’s targets are in opposite directions, assistance is when two targets are identical and cooperation is the case between competition and cooperation. Note that the motion target is defined in a body-fixed coordinate system, i.e. the origin is at the current position. The works analysed in this section are summarized in table 1.

### 4.1. Adaptation of robot target

#### 4.1.1. Competition

One embodiment of competition is when the robot target  $\mathbf{x}^*$  is opposite to the direction of the human target  $\mathbf{x}_h^*$  ( $\xi_0 = \mathbf{0}, \lambda > 1$  in equation (11)). Taking into account human performance, a rehabilitation robot may deliberately challenge the trainee by amplifying their movement errors to improve the outcomes of physical therapy [102]. Furthermore, if human movement or behaviour can be predicted, the robot can plan its motion in time to respond to the situation overseen by the human [47, 103].

#### 4.1.2. Assistance

Many works on human–robot interaction assume that the human has better perception and decision abilities than the robot and so should have a leading role. Using impedance control [104], the robot can be assigned a passive follower role by setting its stiffness to zero. To make the robot proactively follow the human user, the robot target  $\mathbf{x}^*$  can be set as the human target:

$$\mathbf{x}^* = \mathbf{x}_h^*, \quad (12)$$

(corresponding to equation (11) with  $\xi_0 = \mathbf{0}$  and  $\lambda = 0$ ), requiring to estimate  $\mathbf{x}_h^*$  as has been discussed in section 3.1.

The control community uses adaptive control and iterative learning control to estimate  $\mathbf{x}_h^*$  and adapt  $\mathbf{x}^*$  [105–110], while the ML community uses probabilistic models to encode  $\mathbf{x}_h^*$  and decode it for adaptation of  $\mathbf{x}^*$  [57, 60, 111, 112]. These two approaches may have complementary advantages: the modelling in control theory has more intuitive physical meanings and the system performance can be analysed in rigour; while the ML method takes system uncertainties into account, which are inevitable in a human–robot interaction system.

Table 1. Robot control adaptation.

	Assistance ( $\lambda = 0$ )	Cooperation ( $0 < \lambda < 1$ )	Coactivity ( $\lambda = 1$ )	Competition ( $\lambda > 1$ )
Robot target ( $\xi_0 = \mathbf{0}$ , $\xi_h = \mathbf{x}_h^*$ , $\xi = \mathbf{x}^*$ )	[105–110]: same as human motion target, adaptive/iterative learning control [57, 60, 111, 112]: probabilistic models	[47]: interdependent motion targets	[47]: independent motion targets	[47, 102, 103]: opposite motion target
Robot control command ( $\xi_0 = \mathbf{u}_0$ , $\xi_h = \mathbf{u}_h$ , $\xi = \mathbf{u}$ )	[118, 119, 121, 122]: minimization of human control command	[100, 115–117]: supplement to human control command ( $\mathbf{u}_0 = 0$ ) [118, 119, 121, 122]: interdependent control commands ( $\mathbf{u}_0 \neq 0$ )	[118, 119, 121, 122]: independent control commands	
Robot gain ( $\xi_0 = \mathbf{K}_0$ , $\xi_h = \mathbf{K}_h$ , $\xi = \mathbf{K}$ )	[5, 132–135]: identical to human gain	[125, 127, 128]: interaction stability		[31, 117, 129–131, 134, 143–146]: complementary gain adaptation
Metrics for performance evaluation	Trajectory tracking error, force tracking error, task completion time	Trajectory tracking error, human effort, task completion time, human–robot disagreement	Task completion time	Force tracking error, distance to obstacles, impact force

#### 4.1.3. Cooperation

When the robot and human have different targets, switching between them to pass the leading role has been considered in several works, see [107]. However, very few works have explored the intermediate behaviours between leading (corresponding to Competition or Co-activity) and submission (through the Assistance behaviour). Initial effort has been made in [47] to continuously adapt  $\mathbf{x}^*$  according to  $\mathbf{x}_h^*$  from competition, cooperation to assistance, depending on the task and system's state. There are also works to realize cooperation by adapting the controller  $\mathbf{u}_r$ , as discussed in the following section.

#### 4.2. Adaptation of robot control command

Many early works use this strategy to address safety in robot-assisted surgery. In these works, a robot controller  $\mathbf{u} = \mathbf{u}_0$  (setting  $\lambda = 1$  in equation (11)) is designed based on the knowledge of the task and environment, without necessarily considering the human partner's behaviour. Virtual fixtures in the form of force fields have been proposed to keep away from dangerous areas (with a repulsive force) or guide movement (through an attractive force) [20, 113, 114].

In equation (1),  $\mathbf{u}$  can be adapted with  $\mathbf{u}_h$  to modulate the interaction between the robot and human. For instance, to increase the human user's loading capability or reduce their effort, an exoskeleton can provide an assistive force or torque with (setting  $\mathbf{u}_0 = \mathbf{0}$  and  $1 > \lambda > 0$  in equation (11)):

$$\mathbf{u} = (1 - \lambda) \mathbf{u}_h, \quad (13)$$

can be modulated according the context of a task [100]. Dynamic role allocation can be thus be implemented by adapting  $\mathbf{u}$  depending on  $\mathbf{u}_h$ . In [115] where a robot and a human move a table to a target position without a specified path, it is shown that an adaptive role strategy is superior to a constant role strategy in terms of reduced human effort and task completion time, although human users prefer the constant role. While no explicit relationship between  $\mathbf{u}$  and  $\mathbf{u}_h$  is predefined, adaptation of  $\mathbf{u}$  according to  $\mathbf{u}_h$  is studied for a collaborative search and rescue task [116]. Similar to that in [117], the robot tries to reach its own target when there is no human input and gives up its leading role to the human partner when they intervene.

In systems where it is possible to evaluate the confidence level of the robot  $0 \leq \alpha \leq 1$  in performing a task, a higher-order logic can be used [118] to modulate the command between an active component (corresponding to the known part) and a passive component (for the unknown part), yielding the nominal controller equation (11):

$$\mathbf{u}_0 = \alpha \mathbf{u}_a + (1 - \alpha) \mathbf{u}_p. \quad (14)$$

Typically, the active part  $\mathbf{u}_a$  may be implemented as elastic guidance along a planned trajectory and the passive part  $\mathbf{u}_p$  as damping yielding to caster dynamics and stabilising the system. Using the above strategy, haptic assistance is investigated for tele-operation in [119], which mimics the human decision-making process and improves performance compared to no assistance [120].

A finite-state machine based role exchange (RE) mechanism is discussed in [121, 122] in the context of virtual haptic board game, where the dominant role and associated control level are dynamically changed according to rendered stiffness.

### 4.3. Adaptation of robot gain

#### 4.3.1. Interaction stability

Many early literature about adaptation of the robot's gains concerns the interaction stability, which is not surprising as stability is critical to ensure safety for human-in-the-loop systems. Passivity is a widely adopted criterion to guarantee the interaction stability [123], based on which the robot is controlled using an impedance model with fixed robot gains [124]. Such a conservative controller makes the robot behave as a passive follower and it generally increases the human physical load. Many works propose adapting the robot control gains to provide the human user assistance and reduce their effort while ensuring interaction stability. One popular approach is to increase the robot's damping at low velocity in order to ensure stability and decrease it at high velocity to lower the robot's hindrance to the human [125, 126]. A similar idea of adapting robot damping is found in [55, 127, 128], but the adaptation is triggered according to the change of the human force and the estimated human arm stiffness. With some knowledge about the human gains  $\mathbf{K}_h$ , the robot gains can be updated to guarantee complementary stability [129]. Using the estimated interaction stiffness (corresponding to human gains), human-robot interaction stability can be investigated using Lyapunov theory [130], then the robot damping parameters can be adapted to ensure stability [131].

In above works, the robot plays a follower role without motion target, reducing the feedback controller in equation (3) to  $\mathbf{v} = \mathbf{B}\dot{\mathbf{x}}$ . This controller with only a damping term makes the robot behave like a 'brake' which dissipates the system energy. While this is useful in ensuring system stability, it increases the human user effort. By having its own motion target, the robot can provide rich interaction behaviours as are needed in various applications, yielding more options about how to regulate the relationship between  $\mathbf{K}$  and  $\mathbf{K}_h$ , as detailed in the following sections.

#### 4.3.2. Assistance

This type of interaction can be described by (setting  $\mathbf{K}_0 = \mathbf{0}$  and  $\lambda = 0$  in equation (11)):

$$\mathbf{K} = \mathbf{K}_h, \quad (15)$$

which indicates that the robot has a similar controller gain as the human. One direct application of this interaction is to transfer the human operator's variable impedance to the robot so that it can duplicate the operator's movement [5, 132, 133]. In collaborative manipulation of an object along the same direction [134, 135], the robot's impedance is set the same as the human partner's so that motion synchronization is achieved and the object load is equally shared between the human and robot.

#### 4.3.3. Competition and complementary control

This type of interaction can be described by (setting  $\lambda = 2$  and  $\xi_0 = \mathbf{K}_0$  in equation (11)):

$$\mathbf{K} = \mathbf{K}_0 - \mathbf{K}_h, \quad (16)$$

where  $\mathbf{K}_0$  is a predefined matrix corresponding to the task specification. If  $\mathbf{K}_0$  overrides  $\mathbf{K}_h$ , the robot will either provide complementary assistance to the human, or compete with them. How should this interaction behaviour be implemented? Let us consider two cases: when the knowledge of a higher-level task objective is predefined, and when it is not.

In physical rehabilitation [136], the task target is usually predefined and the robot is expected to complement the human partner's force required to complete the task. In this case, it is not necessary to explicitly define  $\mathbf{K}_0$  and one can use the error to the task target to indirectly evaluate the human performance. Therefore, the robot's gains or impedance parameters can be increased when the task cannot be achieved by the human and decreased otherwise [137–139]. This control strategy is closely related to adaptive control [140] and iterative learning control [138, 141, 142]. It is also possible to predefine  $\mathbf{K}_0$  as a controller gain that can achieve the known task target [31]. Then, in an extreme case where the human user's gain  $\mathbf{K}_h = \mathbf{0}$ , according to equation (16) the robot will provide sufficient actuation with  $\mathbf{K} = \mathbf{K}_0$  to complete the task; in the dual case where  $\mathbf{K}_h = \mathbf{K}_0$ , the robot will not provide any actuation with  $\mathbf{K} = \mathbf{0}$ ; in a case where  $\mathbf{K}_h > \mathbf{K}_0$ ,

the robot will resist the human movement with  $\mathbf{K} < \mathbf{0}$ . In this rehabilitation application, the motion targets of the robot and human are the same as the task target, i.e.  $\mathbf{x}_t = \mathbf{x}^* = \mathbf{x}_h^*$ . In the case where  $\mathbf{x}^* \neq \mathbf{x}_h^*$ , researchers used the criterion in equation (16) to adapt the robot's leading role in various applications. In human–robot handshaking [143], the authors used a hidden Markov model to determine if the human partner is in passive or active state based on the human limb parameters: if the human partner is passive, then the robot gain is set at a high value to follow its own reference trajectory; otherwise at a low value to make the robot follow the human lead. In a collaborative sawing task [134], the human partner's arm stiffness is estimated using EMG signal and the robot's stiffness was adapted according to equation (16) to provide complementary assistance.

When the knowledge of a higher-level task objective is not available, the decision making for the robot may become tricky. In co-manipulation tasks [117, 144–146], the robot impedance is adjusted to a high value in a routine task without human's intervention, and automatically decreases when the human partner wants to take over the lead for fine tuning of the robot's path or dealing with unexpected situations, e.g. obstacle avoidance. This means that the robot gives up leading to the human when there is a conflict. This is based on a hypothesis that the human's decision should prevail, which may be questionable in some cases. For example, in shared driving of a car [147] or of a wheelchair [148], the robot should take over the leading role when it perceives an obstacle that the human user does not see. Note that a competition behaviour is inherently used when the robot and human have different motion targets, but enhancing the competition by adapting  $\mathbf{K}$  according to the human parameters has not been studied in the literature, which we believe is worth further investigation.

#### 4.3.4. Adaptation of $\mathbf{K}$ according to $\mathbf{x}_h^*$ or $\mathbf{u}_h$

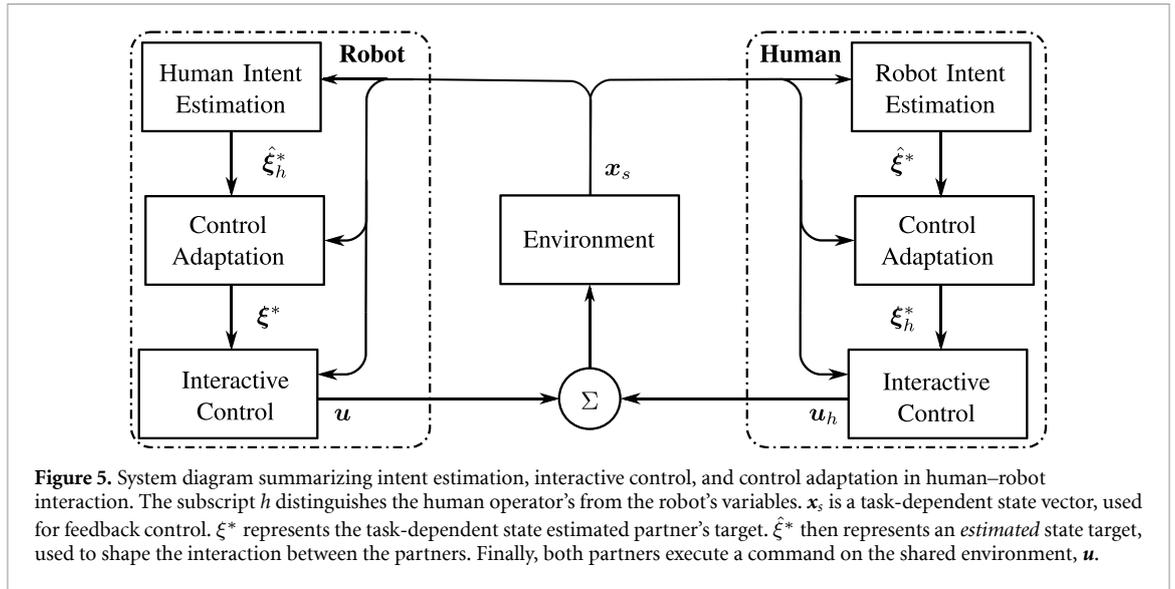
Other works do not prescribe a relation with  $\mathbf{K}_h$ , but adapt  $\mathbf{K}$  in a heuristic ways related to  $\mathbf{x}_h^*$  or  $\mathbf{u}_h$ . For example, in [55], the robot's damping parameter is adapted according to the derivative of the human's applied force to make the interaction more transparent. In [149, 150], a probabilistic task model is established through learning by demonstration and used by the robot to predict the human behaviour in a collaborative task. The human force is used as an indicator of the disagreement between the predicted and actual human behaviors and to adapt the robot's controller gains to either follow its own reference trajectory or be compliant to the human partner. In [151], by estimating the human user's intentional walking direction, the cane robot's mass and damping parameters are adapted with less resistance to the human user in the walking direction and more in the perpendicular direction.

## 5. Interactive control

The two previous sections have reported techniques to adapt the control of contact robots to the human user in a unilateral way, based on information about their motion intent gained from haptic sensing. However, there is ample experimental evidence that humans automatically adapt to their dynamic environment [26, 40, 152]. This means that the user of a contact robot will likely also adapt to its interaction. It has been further observed that when humans carry out connected actions such as transporting and manipulating objects together or in tango dancing, they continuously exchange sensory information to adapt their control and improve their performance [14–16, 153]. How to consider these human motor adaptations? Could contact robots interact and co-adapt with their user as interacting humans do? The techniques that have been developed in recent years to address these questions will be presented in the next three subsections. Figure 5 summarizes the sensorimotor interaction between the human and robot agents, specifying how motion intention and adaptation can be used to develop (bilateral) interactive control.

### 5.1. Sensory augmentation

Humans can combine haptic information from the interaction with another human or with a robot carrying out the same task to improve the accuracy of their own visual sensing [15]. Section 3 has presented techniques enabling a contact robot to detect the motion intent of its user from the interaction force, the common position or other haptic information. The robot can combine this haptic information with information from its other sensors such as camera, ultrasonic, radio-frequency identification. This can improve performance and in turn provide more accurate information to the human user. This interactive control strategy was implemented on a tracking task carried out with a human and a robot equipped with position and force sensors, where it exhibited increased accuracy with the additional interaction information [19]. Study of performance on groups of two, three and four connected humans [16] suggests that the benefit of this sensory augmentation approach increases with the group size, which may transfer to groups of humans and robots interacting with each other.



## 5.2. The game theory of interactive control

How should two interactive agents such as a contact robot and its human user, both able to read the other’s motion intent, share the effort to carry out a common task? Differential GT [154] is a mathematical construct that can be used to investigate the control of these two interacting agents and yield appropriate control. While the human operator may wish to operate the robot according to their planned movement, the robot controller should be able to detect a novice or tired operator and compensate for their shortcomings. Mathematically, the robot will try to minimize its error and effort, as well as the estimated error and effort of the human, computed from a model of the human motor control:

$$C = \int_0^T (e' Q e + u' R u) dt, \quad (17)$$

$$C_h = \int_0^T (e_h' Q_h e_h + u_h' R_h u_h) dt,$$

where  $Q$ ,  $Q_h$ ,  $R$ ,  $R_h$  are matrices weighting the error to effort costs. The constraint conditions for minimizing equation (17) are subject to equation (1). This system of equations from non-cooperative GT determines the *Nash equilibrium*. Using the estimated intent and gain estimated from the human side, the robot can consider the human control by further regulating its control gains.

Following this approach, [155, 156] discuss the strategy design to overcome conflicting intents between human and autonomous vehicle. GT-based controllers for shared driving have been developed in [2, 157, 158], where however the human driver’s controller is assumed to be known. A method to identify the human partner’s control during movement and compute the Nash equilibrium is developed in [31, 159]. This results in improved stability and reduced gains relative to co-activity implemented with one LQR controller for each of the human and robot [31]. Inverse optimal control [160, 161] and Q-learning RL [162] have also been used to identify the human’s cost function for a game theoretic robot controller.

The Nash equilibrium models the situation where the human and robot cannot trust each other. When the operator gets accustomed to the robot’s behaviour and knows that it can assist them, the interaction may correspond more to the *Stackelberg equilibrium* [163, 164] in which the human becomes dominant. Correspondingly the human cost  $C_h$  is first minimized, yielding a solution which is then used to minimize the robot cost  $C$ . This has been applied to model human driving control in [165, 166]. In situations where the human and robot consider both of their tasks simultaneously, this cooperative game can be modelled through the *Pareto equilibrium* where effort tradeoffs both the human and robot tracking errors. This was applied to an object transportation task in [167].

## 5.3. Flexible interaction behaviour through negotiation theory

The GT concept has been used in [101] to specify and classify the human–robot interaction behaviours of figure 4, which have been implemented in [31]. However, the relationship may arguably vary during operation according to changes in the dynamic environment, with human’s fatigue, etc. Therefore, the human–robot relationship should be dynamically modified with a higher level decision process that:

- allows for different perceived information levels and limited communication channels,
- reflects the decision-making and reasoning process explicitly, and
- allows for iterations until the determinate agreement is obtained.

This can be implemented using *negotiation theory* [168, 169] providing an iteration process between two agents considering both intent and effort levels and leading to co-adaptation in three steps. Utility cost functions  $C, C_h$  are first computed to evaluate respective offers  $\mathbf{O}, \mathbf{O}_h$  based on which acceptance will be decided:

$$A[C_h(\mathbf{O}_h), C(\mathbf{O})] = \begin{cases} \text{accept} & C(\mathbf{O}) \geq C_h(\mathbf{O}_h) \\ \text{decline} & C(\mathbf{O}) < C_h(\mathbf{O}_h) \end{cases}. \quad (18)$$

Concession (assistance), tit-for-tat (cooperation and co-activity) and competition are defined in the context of negotiation theory [170] and have been considered to solve collaborative and conflict situations in haptic human–robot interaction [171], which correspond to the cases of  $\lambda = 0$ ,  $0 < \lambda \leq 1$  and  $\lambda > 1$  in figure 4. A feasible opponent model including intent and utility function parameter estimation has been developed based on optimization subject to multiple constraints [172], which does not require prior determination of hypotheses in Bayesian learning [173, 174]. Event-based game is considered to explicitly reflect the negotiation process with complete information of system dynamics in [175], where each player is able to access others' cost functions (i.e. strategies) and then determine the associated strategy. With the support of touchpad-like interface, a butterfly negotiation model allows for multiple goal-directed levels of human–machine interaction [176].

## 6. Limitations and future directions

Contact robots offer many opportunities for interaction with the human user, which have been and are still being extensively investigated. A great deal of progress has been achieved in the development of interaction control systems for contact robots in recent years, as has been reviewed above. In this section we would like to identify a number of open questions and directions to further research and developments.

First, despite abundant effort in the literature, interaction control is still constrained by limited capabilities of intent estimation. Critical to detect motion intent of the human operator is the model of their sensorimotor control. Models should be developed with few parameters that can be identified during movement, while representing the major characteristics of human neuromechanics [26].

While human intent can be represented in different forms, the communicated information in physical human–robot interactions first consists of the interaction force, position, velocity, etc. This sensory information can be combined with the robot's sensing to extract maximum knowledge about the ongoing task as was described under sensory augmentation in section 5.1. Intent estimation is mathematically looking for the single right solution from multiple (or indefinite) possibilities, which is challenging given limited available information. This may be facilitated by using multi-modality interaction integrating other information exchange such as vision, natural language, etc [177].

There are significant gains to be made by the sensing and analysis of biosignals of the human operator to improve the intent estimation. For EMG signals for example, methods are required for handling the diversity of the signals across tasks, the non-stationarity of signals over time and experience, and the variation of signals across human users. High-density EMG coupled with advanced signal processing and ML may provide maximum information about the neural signals in a non-invasive way [178]. Advances would also be provided by practical interfaces, for example wearable interfaces that can be easily donned such as EMG bracelets [179]. Other techniques to acquire signals stemming from muscle activities, such as wearable ultrasound devices, are being developed [180].

It is necessary to develop better task representations for performing and learning interacting in more complex tasks, as discussed in section 3.3. This is not a problem specific to contact robots, and rather corresponds to a diverse range of applications and use cases. A core problem for learning task representations from interaction data is the *causal confusion* challenge whereby purely observational data does not help learn an accurate model of the underlying system [181], and so more explicit models of how observed variables are related to the underlying phenomenon are required to separate cause and effect in observations. Another fundamental issue in learning task representations is that of *covariate shift*, where the nature of the underlying task changes over time and the original model no longer performs well [70]. This has implications on the task-specific nature of learning task representations from demonstrations or observations, where a robot system can be limited to performing one task, or a limited set of tasks based on what it has been shown previously.

Humans are able to generalize learning, for example the learning of a small bicycle as a child automatically transfers to the driving of any larger bicycle when one grows [182]. Similarly, an emerging area of research which may help improve the generality of assistance that can be provided by contact robotics is that of meta-learning, whereby data from an unbounded set of tasks can be incorporated to a general model that can then be rapidly adapted to a new unobserved task [183]. Applying such task learning methods to contact robotics could allow for generally assistive robot systems that can more easily shift from task to task, with minimal re-training or adaptation required.

On the control side, it has been discussed that there is an opportunity to integrate control theory and ML. Among many other reasons, one is that the current control theory for physical human–robot interactions is mainly limited to a single task, e.g. moving an object from one position to another, while ML has been used to deal with multiple tasks. However, one shortcoming of ML is lack of strict analysis of system stability and performance, which is crucial to physical human–robot interactions. It is important to investigate how to combine their complementary advantages.

Regarding the interaction behavior, a lot of effort has been made on cooperation between human and robot, but not explicitly on competition. Therefore, we want to draw attention on the cases where the robot may have alternative objective or criteria for executing the overall goal that conflicts with the human intent. This is because competition is inherent to many applications where it is used to prevent human errors such as virtual fixtures for obstacle avoidance. Versatile physical human–robot interactions may be developed by adapting between assistance and competition according to human input.

As we have seen above, many works in physical human–robot interaction have considered the human partner as a system to identify, without considering their own adaptation to the robot. How will this co-adaptation affect interaction stability? Will a new model considering co-adaptation improve prediction of human behaviours and even guide the modulation of human learning? Works discussed in section 5 have presented preliminary results, but have not addressed adequately how human adaptation can be modelled and integrated to robot adaptation. Conversely, researchers in the field of human motor control mainly focus on the human side. Some results could already be used to design human-like robotic control strategies, but more benefits can be expected by considering human–robot co-adaptation, which is thus worth further study.

Last but not least, there are still relatively few works on how human users perceive the mechanical interaction with a robot. In general, it was found that operators prefer interacting with a predictable behaviour than with human-like control or even with a human [184, 185]. However, this does not mean that human-like behaviours should be avoided. Traditional contact robots have used either trajectory guidance such as in rehabilitation robots or caster dynamics for collaborative tasks such as carrying an object together [186]. However it was shown that this modifies the behaviour and interferes with learning [36], in contrast to human-like interaction control [187]. Furthermore, the potential of distinct interactive control strategies such as proposed in [101] has not been well exploited yet.

## Data availability statement

No new data were created or analysed in this study.

## Acknowledgments

The work was supported in part by the UK EPSRC Grant Nos. EP/T006951/1 and EP/R026092/1; the EC H2020 COST ACTION TD16116, EC H2020 ICT 871803 CONBOTS, FETOPEN 829186 PH-CODING and ICT 871767 REHYB programme grants; and the Slovenian Research Agency P2-0076.

## ORCID iD

Aran Sena  <https://orcid.org/0000-0001-5453-2940>

## References

- [1] Ganesh G, Takagi A, Osu R, Yoshioka T, Kawato M and Burdet E 2014 Two is better than one: physical interactions improve motor performance in humans *Sci. Rep.* **4** 1–7
- [2] Flad M, Otten J, Schwab S and Hohmann S 2014 Steering driver assistance system: a systematic cooperative shared control design approach *IEEE Int. Conf. on Systems, Man and Cybernetics (SMC)* pp 3585–92
- [3] Taylor R H, Menciassi A, Fichtinger G, Fiorini P and Dario P 2016 *Medical Robotics and Computer-Integrated Surgery* (Berlin: Springer) pp 1657–84
- [4] Semprini M, Laffranchi M, Sanguineti V, Avanzino L, De Icco R, De Michieli L and Chiappalone M 2018 Technological approaches for neurorehabilitation: from robotic devices to brain stimulation and beyond *Frontiers Neurol.* **9** 1–9

- [5] Ajoudani A, Zanchettin A M, Ivaldi S, Albu-Schäffer A, Kosuge K and Khatib O 2018 Progress and prospects of the human–robot collaboration *Auton. Robots* **42** 957–75
- [6] Shafiq A, Ataka A, Lazpita B U, Shiva A, Wurdemann H and Althoefer K 2019 Real-time robot-assisted ergonomics *IEEE Int. Conf. on Robotics and Automation (ICRA)* vol 2019 pp 1975–81
- [7] Sherwani F, Asad M M and Ibrahim B 2020 Collaborative robots and industrial revolution 4.0 (IR 4.0) *Int. Conf. on Emerging Trends in Smart Technologies (ICETST)* pp 1–5
- [8] Hesse S, Werner C, Pohl M, Rckriem S, Mehrholz J and Lingnau M L 2005 Computerized arm training improves the motor control of the severely affected arm after stroke: a single-blinded randomized trial in two centers *Stroke* **36** 1960–6
- [9] Kahn L E, Lum P S, Rymer W Z and Reinkensmeyer D J 2006 Robot-assisted movement training for the stroke-impaired arm: does it matter what the robot does? *J. Rehabil. Res. Dev.* **43** 619
- [10] Balasubramanian S, Klein J and Burdet E 2010 Robot-assisted rehabilitation of hand function *Curr. Opin. Neurol.* **23** 661–70
- [11] Burgar C G, Lum P S, Scremin A M, Garber S L, der Loos H F V, Kenney D and Shor P 2011 Robot-assisted upper-limb therapy in acute rehabilitation setting following stroke: department of Veterans Affairs multisite clinical trial *J. Rehabil. Res. Dev.* **48** 445–58
- [12] Colombo R and Sanguinet V (eds) 2018 *Rehabilitation Robotics: Technology and Application* (New York: Academic)
- [13] Colgate J E, Bicchi A and Peshkin M 2008 Safety for physical human–robot interaction *Springer Handbook of Robotics* (Berlin, Heidelberg: Springer) pp 1335–48
- [14] Sowers A and Ting L H 2014 Perspectives on human-human sensorimotor interactions for the design of rehabilitation robots *J. NeuroEng. Rehabil.* **11** 142
- [15] Takagi A, Ganesh G, Yoshioka T, Kawato M and Burdet E 2017 Physically interacting individuals estimate the partner’s goal to enhance their movements *Nat. Human Behav.* **1** 1–6
- [16] Takagi A, Hirashima M, Nozaki D and Burdet E 2019 Individuals physically interacting in a group rapidly coordinate their movement by estimating the collective goal *eLife* **8** 1–19
- [17] Beckers N, van Asseldonk E and Kooij H 2020 Haptic human–human interaction does not improve individual visuomotor adaptation *Sci. Rep.* **10** 1–11
- [18] Takagi A, Beckers N and Burdet E 2016 Motion plan changes predictably in dyadic reaching *PLoS One* **11** 1–15
- [19] Li Y, Eden J, Carboni G and Burdet E 2020 Improving tracking through human–robot sensory augmentation *IEEE Robot. Autom. Lett.* **5** 4399–406
- [20] Xia T, Kapoor A, Kazanzides P and Taylor R 2011 A constrained optimization approach to virtual fixtures for multi-robot collaborative teleoperation *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 639–44
- [21] Díaz I, Gil J J and Sánchez E 2011 Lower-limb robotic rehabilitation: literature review and challenges *J. Robot.* **2011** 1–11
- [22] Ivanova E, Eden J, Zhu S, Carboni G, Yurkewich A and Burdet E 2021 Short time delay does not hinder haptic communication benefits *IEEE Trans. Haptics* **14** 322–7
- [23] Jarrassé N, Sanguineti V and Burdet E 2014 Slaves no longer: review on role assignment for human–robot joint motor action *Adapt. Behav.* **22** 70–82
- [24] Kim W, Peternel L, Lorenzini M, Babič J and Ajoudani A 2021 A human–robot collaboration framework for improving ergonomics during dexterous operation of power tools *Robot. Comput.-Integr. Manuf.* **68** 1–10
- [25] Losey D P, McDonald C G, Battaglia E and O’Malley M K 2018 A review of intent detection, arbitration and communication aspects of shared control for physical human–robot interaction *ASME Appl. Mech. Rev.* **70** 1–19
- [26] Burdet E, Franklin D W and Milner T E 2013 *Human Robotics: Neuromechanics and Motor Control* (Cambridge, MA: MIT Press)
- [27] Carboni G, Nanayakkara T, Takagi A and Burdet E 2021 Adapting the visuo-haptic perception through muscle coactivation *Sci. Rep.* **11** 1–7
- [28] Slotine J-J E and Li W 1991 *Applied Nonlinear Control* (Englewood Cliffs, NJ: Prentice-Hall)
- [29] Jarrassé N, Charalambous T and Burdet E 2013 A framework to describe, analyze and generate interactive motor behaviors *PLoS One* **7** 1–13
- [30] Peshkin M A, Colgate J E, Wannasupphrasit W, Moore C A, Gillespie R B and Akella P 2001 Cobot architecture *IEEE Trans. Robot. Autom.* **17** 377–90
- [31] Li Y, Carboni G, Gonzalez F, Campolo D and Burdet E 2019 Differential game theory for versatile physical human–robot interaction *Nat. Mach. Intell.* **1** 36–43
- [32] Berret B and Jean F 2020 Stochastic optimal open-loop control as a theory of force and impedance planning via muscle co-contraction *PLOS Comput. Biol.* **16** 1–28
- [33] Berret B, Conessa A, Schweighofer N and Burdet E 2021 Stochastic optimal feedforward-feedback control determines timing and variability of arm movements with or without vision *PLOS Comput. Biol.* **17** 1–24
- [34] Wang C, Xiao Y, Burdet E, Gordon J and Schweighofer N 2016 The duration of reaching movement is longer than predicted by minimum variance *J. Neurophysiol.* **116** 2342–5
- [35] Haninger K, Hegeler C and Peternel L 2021 Model predictive control with Gaussian processes for flexible multi-modal physical human robot interaction (arXiv:2110.12433)
- [36] Özen O, Buetler K A and Marchal-Crespo L 2021 Promoting motor variability during robotic assistance enhances motor learning of dynamic tasks *Frontiers Neurosci.* **14** 1–8
- [37] Burdet E, Osu R, Franklin D W, Yoshioka T, Milner T and Kawato M 2000 A method for measuring endpoint stiffness during multi-joint arm movements *J. Biomech.* **33** 1705–9
- [38] Munawar A and Fischer G 2016 A surgical robot teleoperation framework for providing haptic feedback incorporating virtual environment-based guidance *Frontiers Robot. AI* **3** 47
- [39] Dragan A D and Srinivasa S S 2013 A policy-blending formalism for shared control *Int. J. Robot. Res.* **32** 790–805
- [40] Shadmehr R and Mussa-Ivaldi F 1994 Adaptive representation of dynamics during learning of a motor task *J. Neurosci.* **14** 3208–24
- [41] Franklin D W, Liaw G, Milner T E, Osu R, Burdet E and Kawato M 2007 Endpoint stiffness of the arm is directionally tuned to instability in the environment *J. Neurosci.* **27** 7705–16
- [42] Just F, Özen O, Bösch P, Bobrovsky H, Klamroth-Marganska V, Rieni R and Rauter G 2018 Exoskeleton transparency: feed-forward compensation vs. disturbance observer *Automatisierungstechnik* **66** 1014–26
- [43] Maeda Y, Hara T and Arai T 2001 Human–robot cooperative manipulation with motion estimation *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* vol 4 pp 2240–5
- [44] Corteville B, Aertbelien E, Bruyninckx H, Schutter J D and Brussel H V 2007 Human-inspired robot assistant for fast point-to-point movements *IEEE Int. Conf. on Robotics and Automation (ICRA) (Roma, Italy)* pp 3639–44

- [45] Flash T and Hogan N 1985 The coordination of arm movements: an experimentally confirmed mathematical model *J. NeuroSci.* **5** 1688–703
- [46] Iuppariello L, D'addio G, Lanzillo B, Balbi P, Andreozzi E, Improta G, Faiella G and Cesarelli M 2019 A novel approach to estimate the upper limb reaching movement in three-dimensional space *Inform. Med. Unlocked* **15** 1–15
- [47] Takagi A, Li Y and Burdet E 2020 Flexible assimilation of human's target for versatile human–robot physical interaction *IEEE Trans. Haptics* **14** 421–31
- [48] Li Y and Ge S S 2014 Human–robot collaboration based on motion intention estimation *IEEE/ASME Trans. Mechatronics* **19** 1007–14
- [49] Xia P, Hu J and Peng Y 2017 Emg-based estimation of limb movement using deep learning with recurrent convolutional neural networks *Artif. Organs* **42** E67–E77
- [50] Su H, Qi W, Li Z, Chen Z, Ferrigno G and De Momi E 2021 Deep neural network approach in EMG-based force estimation for human–robot interaction *IEEE Trans. Artif. Intell.* **2** 404–12
- [51] Balasubramanian S, Garcia-Cossio E, Birbaumer N, Burdet E and Ramos-Murguialday A 2018 Is EMG a viable alternative to BCI for detecting movement intention in severe stroke? *IEEE Trans. Biomed. Eng.* **65** 2790–7
- [52] Bartolozzi C, Natale L, Nori F and Metta G 2016 Robots with a sense of touch *Nat. Mater.* **15** 921–5
- [53] Bergner F, Dean-Leon E and Cheng G 2016 Event-based signaling for large-scale artificial robotic skin - realization and performance evaluation *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 4918–24
- [54] DelPreto J and Rus D 2019 Sharing the load: human–robot team lifting using muscle activity *IEEE Int. Conf. on Robotics and Automation (ICRA) (Montreal, Canada)* pp 7906–12
- [55] Duchaine V and Gosselin C M 2007 General model of human–robot cooperation using a novel velocity based variable impedance control *2nd Joint EuroHaptics Conf. and Symp. on Haptic Interfaces for Virtual Environment and Teleoperator Systems* pp 446–51
- [56] Kronander K and Billard A 2012 Online learning of varying stiffness through physical human–robot interaction *IEEE Int. Conf. on Robotics and Automation (ICRA)* vol 2012 pp 1842–9
- [57] Billard A, Calinon S, Dillmann R and Schaal S 2008 *Robot Programming by Demonstration* (Berlin: Springer)
- [58] Calinon S 2016 A tutorial on task-parameterized movement learning and retrieval *Intell. Ser. Robot.* **9** 1–29
- [59] Rozo L, Calinon S, Caldwell D G, Jiménez P and Torras C 2016 Learning physical collaborative robot behaviors from human demonstrations *IEEE Trans. Robot.* **32** 513–27
- [60] Rozo L 2019 Interactive trajectory adaptation through force-guided Bayesian optimization *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 7596–603
- [61] Peternel L, Petrič T and Babič J 2018 Robotic assembly solution by human-in-the-loop teaching method based on real-time stiffness modulation *Auton. Robots* **42** 1–17
- [62] Havoutis I and Calinon S 2019 Learning from demonstration for semi-autonomous teleoperation *Auton. Robots* **43** 713–26
- [63] Javdani S, Srinivasa S and Bagnell J 2015 Shared autonomy via hindsight optimization *Robot. Sci. Syst.* **37** 1–24
- [64] Peternel L, Petrič T, Oztop E and Babič J 2014 Teaching robots to cooperate with humans in dynamic manipulation tasks based on multi-modal human-in-the-loop approach *Auton. Robots* **36** 123–36
- [65] Nemeč B, Likar N, Gams A and Ude A 2018 Human robot cooperation with compliance adaptation along the motion trajectory *Auton. Robots* **42** 1023–35
- [66] Schaal S 1996 Learning from demonstration *Int. Conf. on Neural Information Processing Systems (NIPS)* pp 1040–6
- [67] Argall B, Chernova S, Veloso M and Browning B 2009 A survey of robot learning from demonstration *Robot. Autom. Syst.* **57** 469–83
- [68] Kober J, Bagnell J and Peters J 2013 Reinforcement learning in robotics: a survey *Int. J. Robot. Res.* **32** 1238–74
- [69] Deisenroth M P, Neumann G and Peters J 2013 *A Survey on Policy Search for Robotics* vol 2 (Massachusetts, Delf: Foundations and Trends in Robotics) pp 1–2
- [70] Osa T, Pajarinen J, Neumann G, Bagnell J, Abbeel P and Peters J 2018 An algorithmic perspective on imitation learning *Found. Trends Robot.* **7** 1–179
- [71] Huang Y, Silvério J, Rozo L and Caldwell D G 2018 Generalized task-parameterized skill learning *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 5667–474
- [72] Sena A, Michael B and Howard M 2019 Improving task-parameterised movement learning generalisation with frame-weighted trajectory generation *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 4281–7
- [73] Hu S and Kuchenbecker K J 2019 Hierarchical task-parameterized learning from demonstration for collaborative object movement *Appl. Bionics Biomech.* **2019** 9765383
- [74] Vogt D, Stepputtis S, Grehl S, Jung B and Ben Amor H 2017 A system for learning continuous human–robot interactions from human-human demonstrations *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 2882–9
- [75] Schaal S 2006 *Dynamic Movement Primitives—A Framework for Motor Control in Humans and Humanoid Robotics* (Tokyo: Springer) ([https://doi.org/10.1007/4-431-31381-8\\_23](https://doi.org/10.1007/4-431-31381-8_23))
- [76] Prada M, Remazeilles A, Koene A and Endo S 2014 Implementation and experimental validation of dynamic movement primitives for object handover *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems* pp 2146–53
- [77] Huang R, Cheng H, Qiu J and Zhang J 2019 Learning physical human–robot interaction with coupled cooperative primitives for a lower exoskeleton *IEEE Trans. Autom. Sci. Eng.* **16** 1566–74
- [78] Huang Y, Rozo L, Silvério J and Caldwell D G 2019 Kernelized movement primitives *Int. J. Robot. Res.* **38** 833–52
- [79] Abu-Dakka F J, Huang Y, Silvério J and Kyrki V 2021 A probabilistic framework for learning geometry-based robot manipulation skills *Robot. Autom. Syst.* **141** 103761
- [80] Barber D 2012 *Bayesian Reasoning and Machine Learning* (Cambridge: Cambridge University Press)
- [81] Sutton R S and Barto A G 2018 *Reinforcement Learning: An Introduction* (Cambridge, MA: MIT Press)
- [82] Arulkumaran K, Deisenroth M P, Brundage M and Bharath A A 2017 Deep reinforcement learning: a brief survey *IEEE Signal Process. Mag.* **34** 26–38
- [83] Liu R, Nageotte F, Zanne P, de Mathelin M and Dresch-Langley B 2021 Deep reinforcement learning for the control of robotic manipulation: a focussed mini-review *Robotics* **10** 20
- [84] Arora S and Doshi P 2021 A survey of inverse reinforcement learning: challenges, methods and progress *Artif. Intell.* **297** 103500
- [85] Zhang T and Mo H 2021 Reinforcement learning for robot research: a comprehensive review and open issues *Int. J. Adv. Robot. Syst.* **18** 3
- [86] Tobin J, Fong R, Ray A, Schneider J, Zaremba W and Abbeel P 2017 Domain randomization for transferring deep neural networks from simulation to the real world *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 23–30

- [87] Zhao W, Queralta J P and Westerlund T 2020 Sim-to-real transfer in deep reinforcement learning for robotics: a survey *IEEE Symp. Series on Computational Intelligence (SSCI)* pp 737–44
- [88] Romera-Paredes B and Torr P 2015 An embarrassingly simple approach to zero-shot learning *Int. Conf. on Machine Learning* vol 37 pp 2152–61
- [89] Jang E, Irpan A, Khansari M, Kappler D, Ebert F, Lynch C, Levine S and Finn C 2021 BC-z: zero-shot task generalization with robotic imitation learning *Annual Conf. on Robot Learning*
- [90] Ibarz J, Tan J, Finn C, Kalakrishnan M, Pastor P and Levine S 2021 How to train your robot with deep reinforcement learning: lessons we have learned *Int. J. Robot. Res.* **40** 698–721
- [91] Boy E S, Burdet E, Teo C L and Colgate J E 2007 Investigation of motion guidance with scooter cobot and collaborative learning *IEEE Trans. Robot.* **23** 245–55
- [92] Zeng Q, Teo C L, Rebsamen B and Burdet E 2008 Collaborative path planning for a robotic wheelchair *Disability Rehabil.: Assist. Technol.* **3** 315–24
- [93] Tykal M, Montebelli A and Kyrki V 2016 Incrementally assisted kinesthetic teaching for programming by demonstration 2016 *11th ACM/IEEE Int. Conf. on Human–Robot Interaction (HRI)* pp 205–12
- [94] Bajcsy A, Losey D P, O'Malley M and Dragan A 2017 Learning robot objectives from physical human interaction *Conf. on Robot Learning (CoRL)*
- [95] Shafti A, Tjomsland J, Dudley W and Faisal A A 2020 Real-world human–robot collaborative reinforcement learning *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 11161–6
- [96] Sena A and Howard M 2020 Quantifying teaching behavior in robot learning from demonstration *Int. J. Robot. Res.* **39** 54–72
- [97] Hellström T and Bensch S 2018 Modeling interaction for understanding in HRI *Explainable Robotic Systems Workshop at HRI* ([https://doi.org/10.475/123\\_4](https://doi.org/10.475/123_4))
- [98] Hellström T and Bensch S 2018 Understandable robots - what, why and how *Paladyn, J. Behav. Robot.* **9** 110–23
- [99] Arami A, van Asseldonk E, van der Kooij H and Burdet E 2020 A clustering-based approach to identify joint impedance during walking *IEEE Trans. Neural Syst. Rehabil. Eng.* **28** 1808–16
- [100] Evrard P and Kheddar A 2009 Homotopy switching model for dyad haptic interaction in physical collaborative tasks *IEEE Worldhaptics Conf.* pp 45–50
- [101] Jarrassé N, Charalambous T and Burdet E 2012 A framework to describe, analyze and generate interactive motor behaviors *PLoS One* **7** 1–13
- [102] Abdollahi F, Case Lazarro E D, Listenberger M, Kenyon R V, Kovic M, Bogey R A, Hedeker D, Jovanovic B D and Patton J L 2014 Error augmentation enhancing arm recovery in individuals with chronic stroke: a randomized crossover design *Neurorehabil. Neural Repair* **28** 120–8
- [103] Rubagotti M, Taunyazov T, Omarali B and Shintemirov A 2019 Semi-autonomous robot teleoperation with obstacle avoidance via model predictive control *IEEE Robot. Autom. Lett.* **4** 2746–53
- [104] Hogan N 1985 Impedance control: an approach to manipulation *Trans. ASME, J. Dyn. Syst. Meas. Control* **107** 1–24
- [105] Li Y and Ge S S 2016 Force tracking control for motion synchronization in human–robot collaboration *Robotica* **34** 1260–81
- [106] Losey D and O'Malley M 2017 Trajectory deformations from physical human–robot interaction *IEEE Trans. Robot.* **34** 126–38
- [107] Losey D P and O'Malley M K 2019 Learning the correct robot trajectory in real-time from physical human interactions *ACM Trans. Human–Robot Interact.* **9** 1–19
- [108] Xia J, Li Y, Yang L and Huang D 2019 Spatial repetitive learning control for trajectory learning in human–robot collaboration *IEEE Conf. on Decision and Control (CDC) (Nice, France)* pp 5568–73
- [109] Xia J, Huang D, Li Y and Qin N 2020 Iterative learning of human partner's desired trajectory for proactive human–robot collaboration *Int. J. Intell. Robot. Appl.* **4** 229–42
- [110] Xia J, Song C, Huang D, Xing X, Ma L and Li Y 2021 Waypoints updating based on ADAM and ILC for path learning in physical human–robot interaction *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 3359–65
- [111] Medina J R, Lawitzky M, Mörtl A, Lee D and Hirche S 2011 An experience-driven robotic assistant acquiring human knowledge to improve haptic cooperation *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 2416–22
- [112] Medina J R, Lawitzky M, Molin A and Hirche S 2013 Dynamic strategy selection for physical robotic assistance in partially known tasks *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 1180–6
- [113] Tee K P, Yan R and Li H 2010 Adaptive admittance control of a robot manipulator under task space constraint *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 5181–6
- [114] Bowyer S A, Davies B L and Baena F R y 2014 Active constraints/virtual fixtures: a survey *IEEE Trans. Robot.* **30** 138–57
- [115] Mörtl A, Lawitzky M, Kucukyilmaz A, Sezgin M, Basdogan C and Hirche S 2012 The role of roles: physical cooperation between humans and robots *Int. J. Robot. Res.* **31** 1656–74
- [116] Chipalkatty R, Droge G and Egerstedt M B 2013 Less is more: mixed-initiative model-predictive control with human inputs *IEEE Trans. Robot.* **29** 695–703
- [117] Li Y, Tee K P, Chan W L, Yan R, Chua Y and Limbu D K 2015 Continuous role adaptation for human–robot shared control *IEEE Trans. Robot.* **31** 672–81
- [118] Thobbi A, Gu Y and Sheng W 2011 Using human motion estimation for human–robot cooperative manipulation *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 2873–8
- [119] Corredor J, Sofrony J and Peer A 2017 Decision-making model for adaptive impedance control of teleoperation systems *IEEE Trans. Haptics* **10** 5–16
- [120] Egelman D M, Person C and Montague P R 1998 A computational role for dopamine delivery in human decision-making *J. Cogn. Neurosci.* **10** 623–30
- [121] Kucukyilmaz A, Sezgin T M and Basdogan C 2013 Intention recognition for dynamic role exchange in haptic collaboration *IEEE Trans. Haptics* **6** 58–68
- [122] Oguz S O, Kucukyilmaz A, Sezgin T M and Basdogan C 2010 Haptic negotiation and role exchange for collaboration in virtual environments *IEEE Haptics Symp.* pp 371–8
- [123] Colgate J E and Hogan N 1988 Robust control of dynamically interacting systems *Int. J. Control* **48** 65–88
- [124] Hogan N 1985 Impedance control: an approach to manipulation-part I: Theory; part II: Implementation; part III: Applications *J. Dyn. Syst. Meas. Control* **107** 1–24
- [125] Ikeura R and Inooka H 1995 Variable impedance control of a robot for cooperation with a human *IEEE Int. Conf. on Robotics and Automation (ICRA)* p 3097

- [126] Ficuciello F, Villani L and Siciliano B 2015 Variable impedance control of redundant manipulators for intuitive human–robot physical interaction *IEEE Trans. Robot.* **31** 850–63
- [127] Tsumugiwa T, Yokogawa R and Hara K 2001 Variable impedance control with regard to working process for man-machine cooperation-work system *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS) (Maui, Hawaii)* pp 1564–9
- [128] Tsumugiwa T, Yokogawa R and Hara K 2002 Variable impedance control based on estimation of human arm stiffness for human–robot cooperative calligraphic task *IEEE Int. Conf. on Robotics and Automation (ICRA)* vol 1 pp 644–50
- [129] Buerger S P and Hogan N 2007 Complementary stability and loop shaping for improved human–robot interaction *IEEE Trans. Robot.* **23** 232–44
- [130] Duchaine V and Gosselin C 2008 Investigation of human–robot interaction stability using Lyapunov theory *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 2189–94
- [131] Hahnel D, Triebel R, Burgard W and Thrun S 2009 Safe, stable and intuitive control for physical human–robot interaction *IEEE Int. Conf. on Robotics and Automation (ICRA) (Kobe, Japan)* pp 3676–81
- [132] Laghi M, Ajoudani A, Catalano M G and Bicchi A 2020 Unifying bilateral teleoperation and tele-impedance for enhanced user experience *Int. J. Robot. Res.* **39** 514–39
- [133] Zeng C, Yang C, Cheng H, Li Y and Dai S L 2021 Simultaneously encoding movement and semg-based stiffness for robotic skill learning *IEEE Trans. Ind. Inform.* **17** 1244–52
- [134] Peternel L, Tsagarakis N and Ajoudani A 2017 A human–robot co-manipulation approach based on human sensorimotor information *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** 811–22
- [135] Yu X, He W, Li Y, Xue C, Li J, Zou J and Yang C 2019 Bayesian estimation of human impedance and motion intention for human–robot collaboration *IEEE Trans. Cybern.* **51** 1–13
- [136] Pehlivan A U, Losey D P and O'Malley M K 2016 Minimal assist-as-needed controller for upper limb robotic rehabilitation *IEEE Trans. Robot.* **32** 113–24
- [137] Emken J L, Benitez R, Sideris A, Bobrow J E and Reinkensmeyer D J 2007 Motor adaptation as a greedy optimization of error and effort *J. Neurophysiol.* **97** 3997–4006
- [138] Yang C, Ganesh G, Haddadin S, Parusel S, Albu-Schaeffer A and Burdet E 2011 Human-like adaptation of force and impedance in stable and unstable interactions *IEEE Trans. Robot.* **27** 918–30
- [139] Li X, Liu Y-H and Yu H 2018 Iterative learning impedance control for rehabilitation robots driven by series elastic actuators *Automatica* **90** 1–7
- [140] Slotine J-J E and Li W 1987 On the adaptive control of robotic manipulators *Int. J. Robot. Res.* **6** 49–59
- [141] Arimoto S 1990 Learning control theory for robotic motion *Int. J. Adapt. Control Signal Process.* **4** 543–64
- [142] Li Y, Ganesh G, Jarrassé N, Haddadin S, Albu-Schaeffer A and Burdet E 2018 Force, impedance and trajectory learning for contact tooling and haptic identification *IEEE Trans. Robot.* **34** 1170–82
- [143] Wang Z, Peer A and Buss M 2009 An HMM approach to realistic haptic human–robot interaction *IEEE Worldhaptics Conf.* pp 374–9
- [144] Li Y, Tee K P, Yan R, Chan W L and Wu Y 2016 A framework of human–robot coordination based on game theory and policy iteration *IEEE Trans. Robot.* **32** 1408–18
- [145] Talignani Landi C, Ferraguti F, Fantuzzi C and Secchi C 2018 A passivity-based strategy for coaching in human–robot interaction *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 3279–84
- [146] Samuel C M T and Tee K P 2019 Unified human–robot shared control with application to haptic telemanipulation *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 2221–6
- [147] Li R, Li Y, Li S E, Zhang C, Burdet E and Cheng B 2020 Indirect shared control for cooperative driving between driver and automation in steer-by-wire vehicles *IEEE Trans. Intell. Transp. Syst.* **22** 1–11
- [148] Zeng Q, Teo C L, Rebsamen B and Burdet E 2008 A collaborative wheelchair system *IEEE Trans. Neural Syst. Rehabil. Eng.* **16** 161–70
- [149] Medina J R, Lorenz T, Lee D and Hirche S 2012 Disagreement-aware physical assistance through risk-sensitive optimal feedback control *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 3639–45
- [150] Medina J R, Lee D and Hirche S 2012 Risk-sensitive optimal feedback control for haptic assistance *IEEE Int. Conf. on Robotics and Automation (ICRA)* pp 1025–31
- [151] Wakita K, Huang J, Di P, Sekiyama K and Fukuda T 2013 Human-walking-intention-based motion control of an omnidirectional-type cane robot *IEEE/ASME Trans. Mechatronics* **18** 285–96
- [152] Burdet E, Osu R, Franklin D W, Milner T E and Kawato M 2001 The central nervous system stabilizes unstable dynamics by learning optimal impedance *Nature* **414** 446–9
- [153] Takagi A, Usai F, Ganesh G, Sanguineti V and Burdet E 2018 Haptic communication between humans is tuned by the hard or soft mechanics of interaction *PLOS Comput. Biol.* **14** 1–17
- [154] Başar T and Olsder G 1998 *Dynamic Noncooperative Game Theory* 2nd edn (Philadelphia: Society for Industrial and Applied Mathematics)
- [155] Anderson S J, Peters S C, Iagnemma K D and Pilutti T E 2009 A unified approach to semi-autonomous control of passenger vehicles in hazard avoidance scenarios *IEEE Int. Conf. on Systems, Man and Cybernetics* pp 2032–7
- [156] Tamaddoni S H, Taheri S and Ahmadian M 2011 Optimal preview game theory approach to vehicle stability controller design *Veh. Syst. Dyn.* **49** 1967–79
- [157] Na X and Cole D J 2013 Linear quadratic game and non-cooperative predictive methods for potential application to modelling driver-AFS interactive steering control *Veh. Syst. Dyn.* **51** 165–98
- [158] Ji X, Yang K, Na X, Lv C and Liu Y 2019 Shared steering torque control for lane change assistance: a stochastic game-theoretic approach *IEEE Trans. Ind. Electron.* **66** 3093–105
- [159] Shen D and Cruz J B 2008 Adaptive state feedback Nash strategies for linear quadratic discrete-time games *IFAC World Congress* vol 41 pp 15433–8
- [160] Rothfuß S, Inga J, Köpf F, Flad M and Hohmann S 2017 Inverse optimal control for identification in non-cooperative differential games *IFAC World Congress* vol 50 pp 14909–15
- [161] Inga J, Bischoff E, Molloy T L, Flad M and Hohmann S 2019 Solution sets for inverse non-cooperative linear-quadratic differential games *IEEE Control Syst. Lett.* **3** 871–6
- [162] Koepf F, Ebbert S, Flad M and Hohmann S 2018 Adaptive dynamic programming for cooperative control with incomplete information *IEEE Int. Conf. on Systems, Man and Cybernetics (SMC)* pp 2632–8

- [163] Abi-Farraj F, Pacchierotti C, Arenz O, Neumann G and Giordano P R 2020 A haptic shared-control architecture for guided multi-target robotic grasping *IEEE Trans. Haptics* **13** 270–85
- [164] Von Stackelberg H 2010 *Market Structure and Equilibrium* (Berlin, Heidelberg: Springer)
- [165] Tian R, Sun L, Bajcsy A, Tomizuka M and Dragan A D 2021 Safety assurances for human–robot interaction via confidence-aware game-theoretic human models (arXiv:2109.14700)
- [166] Schwarting W, Pierson A, Alonso-Mora J, Karaman S and Rus D 2019 Social behavior for autonomous vehicles *Natl Acad. Sci.* **116** 24972–8
- [167] Koh J J, Ding G, Heckman C, Chen L and Roncone A 2020 Cooperative control of mobile robots with stackelberg learning *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)* pp 7985–92
- [168] Rothfuß S, Ayllon C, Flad M and Hohmann S 2020 Adaptive negotiation model for human-machine interaction on decision level *IFAC World Congress* vol 53 pp 10174–81
- [169] Baarslag T, Gerding E H, Aydogan R and Schraefel M C 2015 Optimal negotiation decision functions in time-sensitive domains *IEEE/WIC/ACM Int. Conf. on Web Intelligence and Intelligent Agent Technology (WI-IAT)* vol 2 pp 190–7
- [170] Shell G R 2006 *Bargaining for Advantage: Negotiation Strategies for Reasonable People* (London: Penguin)
- [171] Oguz S O, Kucukyilmaz A, Metin Sezgin T and Basdogan C 2012 Supporting negotiation behavior with haptics-enabled human-computer interfaces *IEEE Trans. Haptics* **5** 274–84
- [172] Rothfuß S, Schmidt R, Flad M and Hohmann S 2019 A concept for human–machine negotiation in advanced driving assistance systems *IEEE Int. Conf. on Systems, Man and Cybernetics (SMC)* pp 3116–23
- [173] Hindriks K and Tykhonov D 2008 Opponent modelling in automated multi-issue negotiation using Bayesian learning *Int. Joint Conf. on Autonomous Agents and Multiagent Systems* vol 1 pp 331–8
- [174] Narayanan V and Jennings N R 2006 Learning to negotiate optimally in non-stationary environments *Int. Workshop on Cooperative Information Agents* pp 288–300
- [175] Rothfuß S, Steinkamp J, Flad M and Hohmann S 2020 Cooperative decision making in cooperative control systems by means of game theory *Frontiers of Dynamic Games* (Springer) pp 245–65
- [176] Rothfuß S, Wörner M, Inga J and Hohmann S 2020 A study on human-machine cooperation on decision level *IEEE Int. Conf. on Systems, Man and Cybernetics (SMC)* pp 2291–8
- [177] Xue T, Wang W, Ma J, Liu W, Pan Z and Han M 2020 Progress and prospects of multimodal fusion methods in physical human–robot interaction: a review *IEEE Sens. J.* **20** 10355–70
- [178] Holobar D and Farina D 2021 Noninvasive neural interfacing with wearable muscle sensors: combining convolutive blind source separation methods and deep learning techniques for neural decoding *IEEE Signal Process. Mag.* **38** 103–18
- [179] 2021 Myo armband (available at: [https://everipedia.org/wiki/lang\\_en/Myo\\_armband](https://everipedia.org/wiki/lang_en/Myo_armband))
- [180] Brausch L, Hewener H and Lukowicz P 2019 Towards a wearable low-cost ultrasound device for classification of muscle activity and muscle fatigue *Int. Symp. on Wearable Computers (ISWC)* pp 20–22
- [181] de Haan P, Jayaraman D and Levine S 2019 Causal confusion in imitation learning *Adv. Neural Inform. Process. Syst.* **32** 11698–709
- [182] Braun D A, Aertsen A, Wolpert D M and Mehring C 2009 Motor task variation induces structural learning *Current Biol.* **19** 352–7
- [183] Harrison J, Sharma A, Finn C and Pavone M 2019 Continuous meta-learning without tasks (arXiv:1912.08866)
- [184] Bussy A 2013 Approche cognitive pour la représentation de l'interaction proximale haptique entre un homme et un humanoïde *PhD Dissertation* Université de Montpellier, France
- [185] Ivanova E, Carboni G, Eden J, Krüger J and Burdet E 2021 For motion assistance humans prefer to rely on a robot rather than on an unpredictable human *IEEE Open J. Eng. Med. Biol.* **1** 133–9
- [186] Bussy A, Gergondet P, Kheddar A, Keith F and Crosnier A 2012 Proactive behavior of a humanoid robot in a haptic transportation task with a human partner *IEEE Int. Symp. on Robot and Human Interactive Communication (ROMAN)* pp 962–7
- [187] Ivanova E, Eden J, Carboni G, Krüger J and Burdet E 2022 Motion tracking with a human vs. robot partner—performance and learning *Scientific Reports* (In Press)