



# Game theory and partner representation in joint action: toward a computational theory of joint agency

Cecilia De Vicariis<sup>1</sup>  · Vinil T. Chackochan<sup>2,3</sup> · Vittorio Sanguineti<sup>1</sup>

Accepted: 22 April 2022

© The Author(s) 2022, corrected publication 2022

## Abstract

The sense of agency – the subjective feeling of being in control of our own actions – is one central aspect of the phenomenology of action. Computational models provided important contributions toward unveiling the mechanisms underlying the sense of agency in individual action. In particular, the sense of agency is believed to be related to the match between the actual and predicted consequences of our own actions (comparator model). In the study of joint action, models are even more necessary to understand the mechanisms underlying the development of coordination strategies and how the subjective experiences of control emerge during the interaction. In a joint action, we not only need to predict the consequences of our own actions; we also need to predict the actions and intentions of our partner, and to integrate these predictions to infer their joint consequences. Understanding our partner and developing mutually satisfactory coordination strategies are key components of joint action and in the development of the sense of joint agency. Here we discuss a computational architecture which addresses the sense of agency during intentional, real-time joint action. We first reformulate previous accounts of the sense of agency in probabilistic terms, as the combination of prior beliefs about the action goals and constraints, and the likelihood of the predicted movement outcomes. To look at the sense of joint agency, we extend classical computational motor control concepts – optimal estimation and optimal control. Regarding estimation, we argue that in joint action the players not only need to predict the consequences of their own actions, but also need to predict partner’s actions and intentions (a ‘partner model’) and to integrate these predictions to infer their joint consequences. As regards action selection, we use differential game theory – in which actions develop in continuous space and time – to formulate the problem of establishing a stable form of coordination and as a natural extension of optimal control to joint action. The resulting model posits two concurrent observer-controller loops, accounting for ‘joint’ and ‘self’ action control. The two observers quantify the likelihoods of being in control alone or jointly. Combined with prior beliefs, they provide weighing signals which are used to modulate the ‘joint’ and ‘self’ motor commands. We argue that these signals can be interpreted as the subjective sense of joint and self agency. We demonstrate

---

Extended author information available on the last page of the article

the model predictions by simulating a sensorimotor interactive task where two players are mechanically coupled and are instructed to perform planar movements to reach a shared final target by crossing two differently located intermediate targets. In particular, we explore the relation between self and joint agency and the information available to each player about their partner. The proposed model provides a coherent picture of the inter-relation of prediction, control, and the sense of agency in a broader range of joint actions.

**Keywords** Game theory · Partner model · State observer · Bayes

## 1 Introduction

Joint action is ubiquitous in activities of daily living. We coordinate our actions with peers most of our time – for example, while carrying objects, dancing, playing sports, making music or during rehabilitation exercises with a therapist. Researchers have started to investigate how humans deal with joint actions, challenging the assumption traditionally held in cognitive psychology that perception, action and higher-level cognitive processes can be understood by investigating individual minds in isolation.

The sense of agency – the subjective feeling of being in control of our own actions and, through them, of events in the external world – has been widely studied in individual action. The sense of agency has been often considered as the outcome of an inference process (Wegner, 2003; Wegner & Wheatley, 1999) which takes place after action completion. Others consider the sense of agency as an immediate component of the (ongoing) perceptual and sensory registration (Blakemore et al., 2002; Frith et al., 2000; Haggard, 2005). Processes which precede the action itself (Chambon et al., 2014) may constitute a ‘prospective’ component of the sense of agency.

During the last decade, a comprehensive theoretical and modeling framework has been developed, which attempts to capture the complexity and the multifaceted nature of the sense of agency (J. W. Moore & Fletcher, 2012; Pacherie, 2008; Synofzik et al., 2008; Zalla & Sperduti, 2015). These studies suggest that the sense of self agency results from cues related to the sensorimotor level (efference copy, sensory feedback, sensory predictions) and cues related to higher-order cognitive processes (intentions, beliefs), and relies on one core idea: the principle of congruence – i.e. the match between prior knowledge and intentions, predictions and actual observations.

Bayesian concepts have been implicated in the formation of the sense of agency (Fletcher & Frith, 2009; Izawa et al., 2016; J. W. Moore & Fletcher, 2012; Synofzik et al., 2013).

Comparatively less attention has been addressed to the sense of agency in joint action. A joint action is any activity in which two or more persons need to coordinate their actions or movements in space and time to generate a change in the external world (Sebanz et al., 2006). The sense of joint agency

can be regarded as the sense of shared control over a joint action and its consequences on the environment. Pacherie (2012, 2014) argued that in joint action agents need to predict not only their own actions' outcomes, but also the consequences of the partner's actions. The development of the sense of joint agency likely depends on matching such joint predictions with the actual action outcomes. While information about ourselves is easily accessible, information about the partner typically has a greater degree of uncertainty, leading to possibly incomplete or ambiguous representations. However, little is known about how we perceive other people and to what extent we account for information about our partner during action production and how this contribute to the sense of joint agency.

Coordination puts additional requirements in action planning and execution with respect to individual action, thus likely affecting the sense of joint agency. Emergent and planned coordination are regarded as facilitators for the development of the sense of joint agency (Knoblich et al., 2011). Roles distribution and the number of persons involved in the action affect the capability of predicting partner's actions in a joint action (Pacherie, 2014). It is not always straightforward to discriminate contributions in an ongoing action, thus affecting both self and joint agency (Dewey et al., 2014; van der Wel, 2015).

Self and joint agency seem to be modulated by different cues (Le Bars et al., 2020), but they coexist in certain situations (Bolt et al., 2016; Bolt & Loehr, 2017; Pacherie, 2012). In some forms of joint action, the players may experience a strong sense of joint agency and a weak sense of self agency. This occurs when players perform similar actions with similar effects and synchronous timing, like marching soldiers (Bolt et al., 2016). In other situations, a strong sense of joint agency is accompanied by an equally strong sense of self agency. This is the case when the players are required to produce coordinated yet distinct and complementary actions. To reconcile these findings, it has been suggested that the sense of joint agency comes in two forms, namely *we-agency* – in which joint agency grows at the expenses of self agency (Gallotti & Frith, 2013; van der Wel, 2015) – and *shared-agency* – in which both the senses of joint and self agency remain high (Bolt & Loehr, 2017; Dewey et al., 2014; Le Bars et al., 2020). At the very least, these observations suggest that we are capable to assess the senses of joint and self agency separately.

Here we propose a general modeling framework to describe how the sense of joint agency emerges from our experience of the goals, ongoing performance and final outcome of our own actions. We take a probabilistic (Bayesian) perspective, which naturally unifies the prospective and retrospective components of agency.

We then extend this model to account for joint actions, by also accounting for observations that we maintain distinct senses of self and joint agency, whose relationship is determined by the structural aspects of the interaction – e.g. the interaction modality and the type of joint task. To demonstrate the consistency of our formulation and to highlight its predictions, we use computer simulations to predict the emergence and evolution of the sense of joint agency in a recently reported sensorimotor interaction scenario (Chackochan & Sanguineti, 2019).

## 2 Computational model of joint agency

### 2.1 General framework

Action generation is the end result of two inter-related processes, i.e. movement control and estimation of the state of the body and the external environment.

Movement control is the process of generating motor commands - muscle activations - on the basis of our movement goals and our belief on the current state (position, velocity) of our own body and the external environment (Wolpert, 1997; Wolpert & Miall, 1996).

State estimation – or sensorimotor integration – is the process of combining descending motor commands (efferent copy) and sensory information (sensory reafference) to estimate the evolution of the state of the body (and the external environment).

Both movement control and state estimation can be understood in terms of optimality principles. There is indeed ample evidence that our nervous system exhibits a close-to-optimal performance in both movement control (Todorov & Jordan, 2002) and sensorimotor integration (Wolpert et al., 1995). Hence the combination of optimal control and optimal estimation is the ideal framework for addressing the sense of agency.

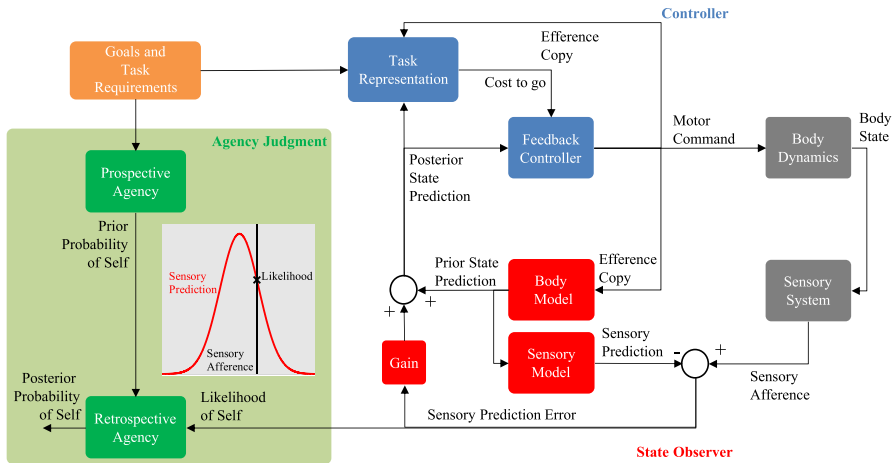
#### 2.1.1 Optimal estimation and agency

Optimal state estimation (Kalman, 1960) is a particular instance of Bayesian estimation – a statistical inference process that optimally combines prior beliefs and the available observations to minimize the prediction uncertainty. In the case of sensorimotor integration (Daniel M. Wolpert et al., 1995), ‘internal models’ maintain representations of the causality between the motor command and the body and environment state – ‘forward models’- and between such state and its sensory consequences – ‘sensory models’ (see Fig. 1). Prior beliefs about body and environment dynamics these internal representations and are combined, within the ‘state observer’, with the mismatch between predicted and actual sensory consequences the motor command - the ‘sensory prediction error’ - to estimate the state of the body and the external world - see Fig. 1.

Theoretical accounts of the sense of agency attribute a crucial role to error signals. In particular, Blakemore et al., 2000 and many others argued that the awareness of being in control of an action reflects the match of the actual and predicted sensory signals – the ‘comparator’ model.

#### 2.1.2 Optimal control and agency

Optimal control posits that humans behave as self-conscious agents, aiming at maximizing their subjective utility, i.e. a trade-off between task-dependent movement cost – in reaching movements, it may simply be the endpoint error – and the perceived



**Fig. 1** Computational model of self agency. The model involves an optimal estimator or state observer block (red) which optimally combines sensory information and motor command to predict the state of the body and the external environment. Here, it is seen as a decision process based on the available information. The agency block (green) combines a prospective or prior component which relies on information on the action goal, and a retrospective or posterior component that reflects the discrepancy between actual and predicted sensory affiance after movement completion. The Optimal controller block (blue) generates the motor commands, which reflect the task requirements, and the current state as predicted by the state observer

effort associated to movement. Hence actions result from a subjective evaluation of their associated costs and/or benefits. This subjective aspect can be summarized by a cost function that incorporates the task goals and requirements.

Given a certain goal, a variety of rules – control policies – can be conceived to map body and environment states into motor commands (Wolpert & Kawato, 1998). There is some evidence that the brain selects the optimal motor command as a trade-off between cost and effort (Shadmehr & Krakauer, 2008; Todorov, 2004) – see Fig. 1.

While the role of optimal estimation has been widely acknowledged in the formation of the sense of agency, optimality of control is at least as important. Awareness of what we are doing and how we are trying to achieve our goals is involved in building the sense of agency (Chambon et al., 2014; J. Moore & Haggard, 2008; Pacherie, 2008). In particular, these multiple cues likely affect our prospective judgement of agency (Lafleur et al., 2020; Synofzik et al., 2013; Zalla & Sperduti, 2015). We may feel to have more or less control over an intended action depending on prior knowledge about the task, the context and the goals to be achieved.

### 2.1.3 A probabilistic perspective

The above considerations point at a comprehensive formulation for the sense of self agency, rooted on probabilistic concepts. Probabilities are intended in subjective or Bayesian terms, i.e. as measures of the degree of belief that some proposition is true.

Specification of the action goals and task requirements and the subjective cost and benefits of an action contribute to the ‘prospective’ component of the sense of agency (Chambon et al., 2014; J. Moore & Haggard, 2008). As the action is initiated, the mismatch between observed and expected sensory outcomes (S. Blakemore et al., 2000; C. D. Frith et al., 2000) combined with the reliability of the sensory afference (J. W. Moore & Fletcher, 2012) continuously modifies such belief. In other words, the sense of agency is a dynamic process, which involves a continuing evaluation of one’s awareness of being in control. At the end of the movement, our overall (‘retrospective’) sense of agency results from the integration of these multiple sources – see Fig. 1.

In probabilistic terms, the prospective and retrospective components of the sense of agency while carrying out an action can be seen as prior and posterior probabilities of being in control. In more formal terms, prospective self agency can be seen as the subjective belief (probability) of being in control:

$$\text{prospective agency (self)} \triangleq \text{Pr (self)}$$

before action onset. Similarly, retrospective self agency can be seen as the probability of being in control after the action is made, conditioned to all evidence collected during action (i.e. the sensory observations, i.e. the sensory inputs observed over the whole action, between time=0 and time=T. We denote sensory observations as  $y(0:T)$ ):

$$\text{retrospective agency(self)} \triangleq \text{Pr (self} \mid y(0 : T))$$

Bayes’ theorem states that posterior and prior probability are related through the ‘likelihood’ of the prediction, which reflects how likely those specific sensory observations are if we assume that we are actually in control. This is quantified as the probability density function of the observation, under the ‘self’ assumption:

$$\text{likelihood(self)} \triangleq p(y(0 : T) \mid \text{self})$$

Likelihood naturally emerges from the inference process underlying state estimation, but is not simply reflected in the mismatch between observed and predicted consequences of action, but more generally in the likelihood of that prediction – a quantity which, besides the mismatch, also reflects the reliability of observations– see Supplementary Materials. Overall, these quantities are related (Bayes theorem) as:

$$\text{retrospective agency(self)} = \frac{\text{likelihood(self)} \times \text{prospective agency(self)}}{p(y(0 : T))}$$

It should be noted that the likelihood accumulates over time as the action proceeds, which suggests that the sense of agency is also a dynamic process. At movement start, the sense of agency is coincident with prospective agency before the action takes place. While the action proceeds, the sense of agency may be at times weaker or stronger.

This formulation naturally extends the comparator model, and in particular implies that the sense of agency is not only strengthened by a greater match, but also by a greater reliability of the sensory information; see also (Izawa et al., 2016; J. W. Moore & Fletcher, 2012). Figure 1 (green block) depicts the relation between likelihood, retrospective and prospective agency.

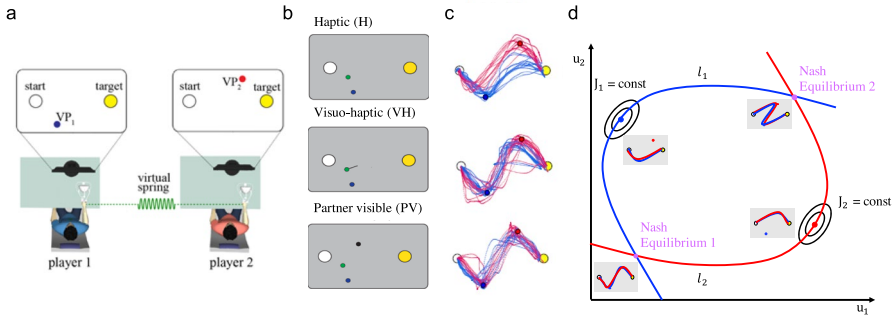
## 2.2 From self agency to joint agency

The above framework can be extended to the sense of joint agency, intended as the subjective sense of shared control over a joint action. We can safely assume that all players are equipped with their own state observer, movement controller – hence the ability to generate their own action - and their subjective sense of self agency.

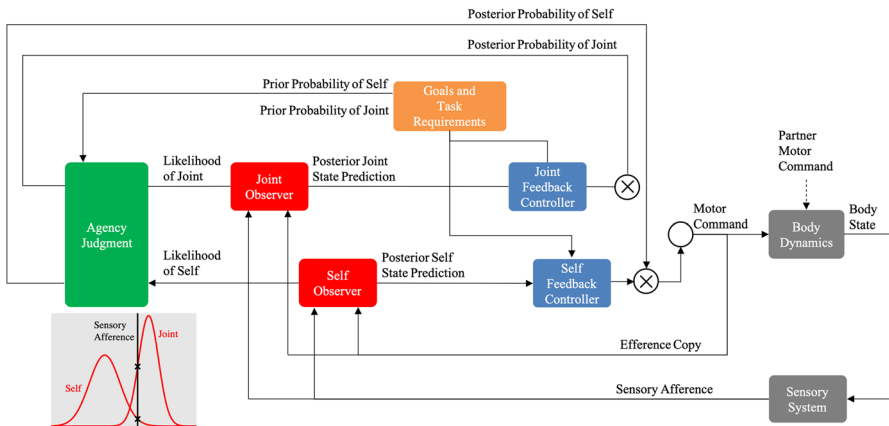
In joint action each player has his/her own goals. The goals may be the same for all players; may be different, or even opposite. However, joint action implies that there is some incentive to act together, which may be a specific task requirement, e.g. the need to reach a shared goal at the same time - or is implied by mechanical coupling e.g. when dancing the waltz. In all cases, acting together always requires that each player accounts for their partner when selecting their action. Ganesh et al. (2014) examined a scenario in which two players are mechanically connected through an elastic force (a virtual spring) and are instructed to track the same moving target. The players share the same goal and, as they are physically coupled, goal fulfillment depends on both self and partner state. Chackochan and Sanguinetti (2019) also focused on a scenario in which two players are mechanically connected. Both were instructed to perform reaching movements through different via-points, while at the same time keeping the interaction forces low (Fig. 2). Again, there is physical interaction and, although the goals are distinct, they still require players' coordination. Bolt et al. (2016) focused on a situation in which two players emit sounds by pressing keys, in alternation according to different rhythmic patterns. In this case there is no physical interaction, but the players need to account for each other in order to synchronize. All the above examples suggest that in joint action scenarios the control policy must necessarily account for both self and partner's state.

### 2.2.1 The partner model

Computational accounts of joint action are relatively rare, and even less (Chackochan & Sanguinetti, 2019; Li et al., 2019; Pesquita et al., 2018; Takagi et al., 2018) explicitly address the need to account for both self and partner state. If there is no mechanical coupling, this can be achieved with distinct 'self' and 'other' state observers; see for instance (Pesquita et al., 2018). However, if two players interact physically, i.e. they exchange forces, their body states cannot be estimated independently. Consequently, to provide reliable estimates the 'self' and 'other' state observers must account for each other. In the words of Noy et al. (2017), 'coupled forward models are necessary for producing co-confident motion'.

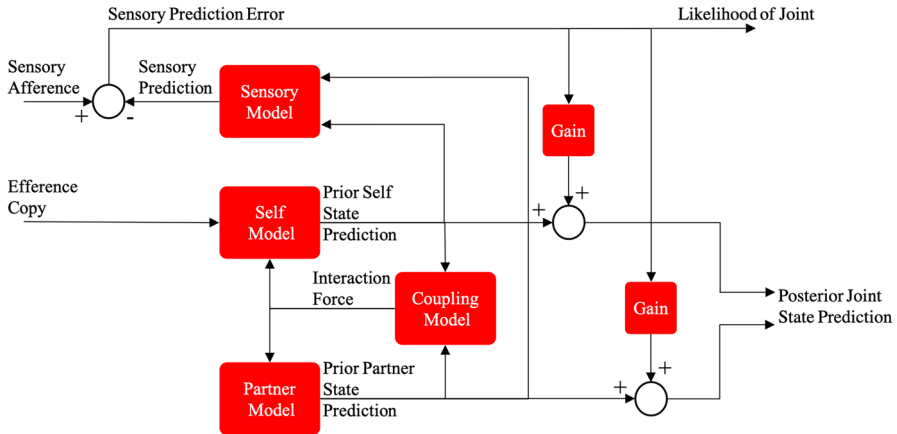


**Fig. 2** A joint action scenario and its game theoretic interpretation. **a)** Experimental apparatus and task. Two players are connected via a virtual spring. They were instructed to perform planar reaching movement from the same initial point to a final target, by crossing different via-points. **b)** Experimental conditions. Three experimental groups different in the amount of information provided about their partner: in the haptic condition (H) players could only feel their partner haptically, through the interaction force; in the visuo-haptic condition (VH), the interaction force vector was displayed on the computer screen; in the partner visible condition (PV), players could see their partner’s ongoing movements through a second cursor displayed on the screen. **c)** Experimental results. Last ten movements of the two players after a training session. **d)** Game Theory predictions. Players have distinct sets of best responses to partner actions, specified by their cost function minima ( $J_1$  and  $J_2$ , displayed by their contours) in the space of possible actions (player 1:  $u_1$ ; player 2:  $u_2$ ). The game has two Nash equilibria – both players passing through both via-points. The Nash equilibria are specified by the intersection of the players’ reaction curves – i.e. the locus of the optimal actions which a player may take for any given action chosen by the partner. The experimental results suggest that as the available information increases (PV group), the players tend to converge to the lowest-effort Nash equilibrium. Modified from Chackochan and Sanguneti (2019).



**Fig. 3** Computational model of joint agency. The assumption that agency is a decision process implies the choice between a set of options. The model involves two optimal estimator (state observer) and optimal controller pairs running in parallel. Each state observer estimates the likelihood of, respectively, self and joint. These quantities are combined with prior information from the task representations to get posterior probabilities of self and joint. These quantities gate the outputs of the optimal controllers and therefore affect the control policy





**Fig. 4** Joint state observer. In addition to a model of own body dynamics (self model), it also maintains a representation of the partner and of the self-partner mechanical coupling (if any). The observer predicts the state of both self and other. These quantities are combined to predict the sensory outcomes which are compared with actual sensory afferences. The sensory prediction error quantifies the joint likelihood. Combined with the prior, it generates the joint sense of agency

In conclusion, in joint action we must posit an additional module as part of the state observer, which accounts for the partner’s state and possibly motor actions, which we will refer as the ‘partner model’ – see Figs. 3 and 4. Estimating the partner state is no different from estimating their own state, in the sense that it requires a forward model of partner’s body and the availability of sensory signals – e.g. vision, proprioception, audition, etc. - which provide information about the partner.

From a computational perspective, partner models may take various forms. They may just estimate the ongoing partner movements by combining measures and prior assumptions. More accurate models may additionally account for the partner’s body dynamics, e.g. inertial properties. For instance they may use internal representations of the partner’s body dynamics to infer their underlying motor commands (Chackochan & Sanguinetti, 2019; Gillijns & De Moor, 2007). Further, partner models may be capable of inferring the partner’s control policy, i.e. the mapping between the state of both players and the partner’s motor command (Li et al., 2019). This type of partner model is much more informative, as it provides not only an estimate of past action, but more generally on the partner’s ongoing strategy, which can be seen as a representation of the partner’s ultimate goal or intentions.

There is no agreement in the literature about the type of partner models and on how they are formed. During joint action players may simply estimate the ongoing partner actions (Chackochan & Sanguinetti, 2019). Other studies suggest that players may develop more general partner representations, also accounting for intentions and ultimate goals (Sebanz et al., 2005). Humans are indeed very good at extrapolating higher order information by observing the motion of their peers: not only they can infer intentions or goals, but can also discriminate between actual human movements from movements that have artificial spatio-temporal features, e.g. they are incompatible with human body biomechanics (Zunino et al., 2020).

Our main focus here is on ‘the mutual give and take between two or more individuals involved in social interactions’ (Chris D. Frith & Frith, 2008), but the ‘partner model’ notion implies, more generally, an ability to infer the intentions, desires and beliefs of others – the Theory of Mind.

We are not explicitly addressing the nature of these processes here. We posit that sensory observations are compared with their predictions under a specific hypothesis to form a likelihood. Likelihoods of the different hypotheses are combined with prior knowledge (priors) into posterior probabilities, which mediate the subsequent action generation. These computational modules constitute a basic form of Theory of Mind.

The current partner model formulation focuses on implicit and immediate perceptual processes (Gallagher, 2008). More explicit and aware forms of reasoning about the partner (Carruthers & Smith, 1996; Gallese & Goldman, 1998) may contribute to higher-level prediction of partner actions, which includes intentions (i.e. the control policies) and ultimate goals (i.e. their objective function). Although these mechanisms are not covered in the current description of the model, they are clearly consistent with the overall architecture. Overall, the proposed model points at a Theory of Mind which posits fast perceptual processes and slower, more cognitively demanding processes which contribute to an efficient understanding of the partner’s mental states (Chris D. Frith & Frith, 2008; Gangopadhyay & Miyahara, 2015; Meinhardt-Injac et al., 2018).

In conclusion, joint action requires that the state observer has an additional ‘partner model’, which may be either distinct or interrelated to the ‘self’ state observer – see Figs. 3 and 4. We suggest that the partner model plays a pivotal role in developing and modulating the sense of agency in joint action.

### 2.2.2 Game theory in joint action

Game theory provides the analytic and computational substrate for the decision-making and control processes underlying joint action. A ‘game’ is a situation ‘involving two or more individuals whose interest are neither completely opposed, nor completely coincident’ (J. F. Nash, 1953). Concepts from game theory have been widely applied in the social sciences to understand how multiple agents coordinate their actions. In a joint action, individuals still aim at maximizing their own subjective utility, but the latter also depends on their partner’s state. This introduces additional complexity in the individual action selection mechanisms, and likely plays a role in forming our sense of joint agency.

In some joint action scenarios, agents agree on a common strategy – for instance, through verbal communication - before the action takes place, thus behaving as a collective (Bacharach, 1999; Newton, 2018; Tuomela, 2007). These situations are referred as ‘cooperative games’. We argue that in these situations the sense of joint agency has a mainly ‘prospective’ character as it largely develops before action initiation.

In other scenarios there is no explicit prior agreement on a common strategy. While we know our goals, we may not be equally aware of the goals of our counterpart. In these situations, coordination emerges gradually as each agent

collects information about their partner actions, their outcomes and possibly their ultimate goals, using various mechanisms (Vesper et al., 2017), during continuous or repeated interaction. Situations in which each player autonomously decides the action to take are referred to as ‘non-cooperative games’. Situations in which none of the players can unilaterally improve their benefit are known as Nash equilibria (J. Nash, 1951).

Differential game theory is the natural extension of optimal control to joint action scenarios as it addresses those situations in which actions develop in time (Başar & Olsder, 1999). In motor control scenarios, differential non-cooperative games have been used to model situations in which humans deal with their counterpart without speaking and by communicating just through sensory cues (visual, acoustic or haptic), but they independently determine their actions (Braun et al., 2009; Chackochan & Sanguineti, 2019; Jarrasé et al., 2012; Li et al., 2019). In sensorimotor versions of the prisoner dilemma and of the rope pulling games, Braun et al. (2009) found that the players converge to Nash equilibria. In contrast, when the same task is performed by one single agent using two hands, they converged to a cooperative solution. These results suggest that, if they have perfect information about their partners, two players are capable of establishing stable coordination strategies which correspond to Nash equilibria. When combined with state and partner estimation, differential game theory can be used to study situations characterized by partial information (Harsanyi, 1967) in which players form conjectures about the other player’s actions. In situations where the players have competing goals and need to negotiate a joint strategy, knowledge about the partner is a major determinant of the resulting coordination strategy. In a recent work in which dyads performed a sensorimotor coordination game, Chackochan and Sanguineti (2019) manipulated the information about the partner available to each player, ranging from relatively unreliable haptic information to highly reliable visual information; see Fig. 2. Depending on the available sensory information, they found that the players converged to different coordination strategies. When the information was more reliable, they converged to a Nash equilibrium. When the information uncertainty increased, they converged to a strategy which assumed there were no partner.

Another aspect of coordination strategies, which is highly relevant to joint agency, is that they often develop gradually. Fictitious play (Berger, 2007; Brown, 1951) has been proposed as a general mechanism for the development of a stable coordination (Chackochan & Sanguineti, 2019; Grau-Moya et al., 2013). In repeated trials, subjects gradually refine their partner model and determine their best response given that partner model. This can be interpreted in Bayesian terms – at each trial, the players combine prior beliefs with the available information on the action outcome to improve their next moves. Nash equilibria are absorbing states in the fictitious play learning process (Fudenberg & Levine, 1998) - once you get there, you stay there.

The sense of joint agency likely evolves as the coordination develops. For instance, the sense of joint agency may be reduced when coordination is unstable

and players have difficulty in synchronizing their actions, but may increase gradually as they converge toward a stable coordination.

### 2.2.3 A modular control architecture to reconcile self and joint agency

The above considerations suggest an extension of the comparator model – or rather the probabilistic model - to account for the sense of joint agency. Like self agency, the subjective experience of acting as a group would have a prospective component. The latter can be intended as the subjective evaluation of costs and benefits associated to the development of a coordination, before the latter takes place. Contextual conditions (Lafleur et al., 2020) and personal traits directly affect the subjective evaluation of tasks and goals (Vallacher & Wegner, 1989). Therefore, they both contribute to shaping the sense of agency.

In probabilistic terms this can be captured by a joint agency ‘prior’:

$$\text{prospective agency (joint)} \triangleq \text{Pr (joint)}$$

After action completion, a ‘joint’ observer, involving a self and a partner model, see Fig. 4, evaluates the ‘joint’ likelihood, i.e. the extent to which the presence of multiple players explains the sensory information  $y(0 : T)$ :

$$\text{likelihood(joint)} \triangleq p(y(0 : T) | \text{joint})$$

Finally, the retrospective sense of joint agency results from the combination of prior assumptions (prospective agency) and the match between actual and expected observations (likelihood) – see Supplementary Materials for details:

$$\text{retrospective agency(joint)} = \frac{\text{likelihood(joint)} \times \text{prospective agency(joint)}}{p(y(0 : T))}$$

How are the sense of self and joint agency related? In a joint action context, either the self or the group may be perceived as being in control of an action. As mentioned in the Introduction, empirical findings suggest that we maintain distinct senses of self and joint agency.

To explain these observations, we argue that during joint action, at least two state observers are active at the same time, respectively accounting for the ‘self’ and ‘joint’ scenario. The ‘self’ observer assumes that the player is alone in contributing to the body and environment state. The ‘joint’ observer assumes that one or more partners are involved in the action.

A general structure for the ‘joint’ observer is depicted in Fig. 4. The joint observer extends the ‘self’ observer of Fig. 1 in that it also includes an internal representation of partner dynamics and of their mechanical coupling (if any); see the Supplementary Material for details. The state predictions of self and other are then combined to form a prediction of the sensory outcome. The sensory prediction error is used to quantify the ‘joint’ likelihood. In the general – see Fig. 4 - the ‘self’ and ‘joint’ observers simultaneously estimate the corresponding likelihoods from the same sensory information. Combined with prior beliefs (prospective

component of self and joint agency), they generate separate retrospective self and joint agency beliefs (self and joint posteriors).

We also argue that the subjective awareness of being in control alone or with other partners also has an influence on action selection: we may choose different actions if we believe we are acting alone, or another player is contributing to our action. To account for this, the model posits two controllers: the ‘joint feedback controller’, which determines a motor command under the assumption of coordinating with a partner, and the ‘self feedback controller’ which assumes that the player acts alone. Each controller generates a motor command based on the state predictions made by the corresponding observer. The self and joint posterior probabilities serve as gating signals to select which motor command will be eventually generated.

The overall model is summarized in Fig. 3. The ‘task’ is represented in terms of goals and requirements, including the requirement of coordinating with one or more partners, and possibly prior agreements with them, and can be captured by a cost function which keeps the interaction into account. This is the prior belief that a player has before the action starts – prospective components of agency. Based on the assumptions of either acting alone or jointly, the two state observers generate different state predictions on the basis of the same sensory information.

An Agency Judgement subsystem combines the prior beliefs with the evidence collected during action (self and joint likelihoods), thus providing an ongoing posterior belief of either acting alone or as a group. During movements the agency judgement continuously gates the motor command generated by the ‘self’ and ‘joint’ controllers.

The proposed architecture is inspired by the ‘MODular Selection And Identification for Control’ (MOSAIC) model (Haruno et al., 2001; D. M. Wolpert & Kawato, 1998). Originally proposed to address the modularity of the motor system, the MOSAIC model posits that the motor system uses multiple observer-controller pairs, each dedicated to a certain aspect of the action and characterized by different assumptions on our body and the environment in which the action takes place. This architecture is more parsimonious than using a single enormously complex model of the external world and accounts for experimental findings suggesting that humans are capable of learning multiple internal representations, switching between them, or combining them, based on the context in which the action is occurring to determine an appropriate motor command – see D. M. Wolpert and Kawato (1998) for more details. The model relies on specialized predictors (observers) for each of the hypotheses under consideration. In joint action these hypotheses include acting alone (‘self’), just observing someone else’s action (‘other’), or acting jointly (‘joint’). The ‘self’ and ‘other’ observers require forward models of, respectively, own and other dynamics. A ‘joint’ observer requires both (plus a model of their mechanical coupling if any). The model keeps duplication of resources at a minimum as some of these modules may be shared at the implementation level.

The proposed model posits a bidirectional relationship between action control and the sense of agency. In this view, agency is a dynamic decision process, which continuously evaluates the assumptions underlying each observer pair. Players reason about themselves – their goals, their role in the action, the perceived features of their body and environment, the presence of additional partners or opponents.

Our behavior changes depending on whether we feel to have a joint control over an event, or we feel we are acting individually. This suggests that the sense of agency is not simply the outcome of an inference process, but also plays a role in determining the joint control policy.

The proposed model implies that the sense of self agency and the sense of joint agency are distinct, though not independent. Rather, they reflect parallel mechanisms that make sense of the available information in different ways. The Agency Judgment block uses all the available information to evaluate what is more likely. One key model prediction is that the likelihood of acting jointly is always greater or equal than the likelihood of acting individually. This is a consequence of the fact that if you add a component in a model, its likelihood with reference to a specific set of observations can only be greater or equal, but never be worse than the model without that component. If there is no partner or the partner plays no role in the interaction, the 'joint' observer would provide no better explanation of the sensory observations than an observer which neglects the partner, hence the 'self' and 'joint' likelihoods would be the same. Conversely, if the partner plays any role in the action, this is reflected in the action's sensory consequences. As such, the 'joint' observer would provide a more likely explanation of the sensory observations than its 'self' counterpart.

Hence the 'joint' observer, which also includes the partner model, always provides more accurate predictions than the 'self' observer if a partner is actually present. Conversely, if no partner were present, the two state observers would lead to identical predictions and consequently identical likelihoods. The relationship between self and joint-likelihood may explain the different experiences of joint agency in different interaction scenarios. In 'we-agency' scenarios like marching soldiers, the partner counts a lot. Therefore, the joint-likelihood is much greater than the self-likelihood. In shared-agency scenarios the two players have complementary roles and tend to be less interdependent. Therefore, joint and self-likelihoods tend to be more similar. It should be noted, however, that the overall sense of agency also depends on the priors – the prospective components of agency – which can make the 'self' posterior probability greater than its 'joint' counterpart. This occurs for instance in extreme situations when the two players do not interact at all and therefore no joint action takes place.

Here we only consider the 'self' and 'joint' scenarios because they are the most relevant for self and joint agency, but the model may even include a 'other' scenario (in which the 'self' does not participate in the action). Indeed, the observers associated to each scenario, the likelihoods they generate and their combinations with the priors may be seen as a Bayesian classifier which, at any given time instant and cumulatively at the end of an action, compares posterior probabilities which can be interpreted as the sense of self and joint agency. This model can account for the attribution problems discussed in Frith et al. (2000).

### 3 Joint agency in a sensorimotor interactive task

To understand the implications of the proposed model, we simulated the task and the experimental protocol used in a published study (Chackochan & Sanguineti, 2019), which specifically addressed the way coordination strategies are affected by uncertainty about the partner; see Fig. 2 for a summary of the experimental protocol and main results.

Simulations have a twofold purpose: (i) to test the internal coherence and completeness of the proposed model; and (ii) to explore model predictions as regards the sense of self and joint agency, in different phases of the experiment and different experimental conditions.

#### 3.1 Task

In the actual experiment, two players were instructed to perform planar arm movements with the same start and end points, but different via-points, while mechanically coupled through a virtual spring. In particular, each player operated the handle of a robot manipulandum, whose movements were mapped into the motion of a cursor on a computer screen placed in front of each player. To simulate the virtual spring, the forces generated by the robots at the level of each end-effector were set to be proportional to the distance between the two hand positions. The participants could not see each other and could not speak. They were not explicitly informed about performing a joint task, but were instructed to keep their interaction forces low, which is an incentive to aim at coordination. The experiment involved three groups of dyads, which differed for the amount of information available about their partner. In the haptic condition (H), the players only perceived their partner through the interaction force. In the visuo-haptic condition (VH) the interaction force was also displayed on the screen, as an arrow attached to their hand cursor. In the partner-visible condition (PV), in addition to the haptic force each player could see their partner as a second cursor moving on the screen. The experimental protocol involved a total of 156 trials. The players were initially not connected (baseline phase, 12 trials). Then the virtual spring was turned on (training phase, 120 trials). Then the connection was removed again (after effect phase, 24 trials). The peculiarity of this task is that the players have different sub-goals and need to negotiate a joint strategy, but have no other cue about their partner's intentions or ongoing actions than the interaction force alone (H group) and additional visual feedback in various forms (VH, PV groups).

#### 3.2 Model implementation

We simulated the two-loop model of Fig. 3. The task is completely specified by a pair of quadratic cost functions, one for each player. The cost functions have four terms, respectively accounting for (i) reaching the final target; (ii) keeping the interaction force low (only present during the training phase); (iii) passing through the via-point; and (iv) keeping the mechanical effort low; see the Supplementary



Material file for details. We modelled the players as a single linear dynamical system (with Gaussian process noise) and two separate sensory systems (with Gaussian measurement noise). The H and VH groups only differed in the uncertainty (measurement noise) of the measured interaction force (lower in the VH group, in which the force is visually displayed to the players). In the PV group, the sensory system had an additional term – visual information about partner’s position and velocity which further reduces the uncertainty about the partner movements.

For each player, we posited two parallel observer-controller loops.

The ‘self’ observer assumes that there is no partner. The ‘joint’ observer predicts the body state by additionally accounting for a partner model, which continuously estimates the ongoing partner actions. Both state observers optimally combine prior state predictions with the current sensory afferences to estimate the next state. The ‘joint’ observer included a partner model which additionally predicts the partner’s ongoing motor command. At each time instant, the two observers compute cumulative self and joint likelihoods, which quantify whether the sensory afferences up to that time instant are best predicted by self or joint action. At each time instant, these measures are combined with a prior probability term which reflects the prior knowledge about the task (prospective sense of agency), to form dynamic self and joint agency beliefs (posterior probabilities). We assumed that all players initially have little or no cues about the presence of a partner. Therefore, in all phases and in all simulations, we assumed that the prospective sense of self agency is slightly greater than the corresponding joint agency. In particular, we set  $\text{prospective}(\text{self})=0.6$  and  $\text{prospective}(\text{joint})=0.4$ . This choice is clearly questionable and the prospective component of agency within this task would deserve a study on its own, as personal traits (Vallacher & Wegner, 1989) and external context (Lafleur et al., 2020) likely affect the prospective component of agency.

The ‘self’ and ‘joint’ controllers have an identical structure, resulting from the same cost function. The ‘joint’ controller additionally incorporates the partner motor command estimated on the previous trial, whereas the ‘self’ controller assumes that there is no partner (the partner motor command is identically zero). This implements the fictitious play learning model. At every time instant, both controllers generate motor commands based on the corresponding state observers. The overall motor command is calculated as the sum of the self and joint motor commands, weighed by the corresponding self and joint agency beliefs.

All details of the model implementation, the values of the model parameters and the technical details of the simulation are reported in a Supplementary Material file.

### 3.3 Simulation results

We simulated the evolution of the behavior of three dyads, respectively in the haptic (H), visuo-haptic (VH) and partner visible (PV) group, over the entire three-phase (baseline, training, aftereffect) 156-trial experimental protocol.



### 3.3.1 Effect of training and partner information

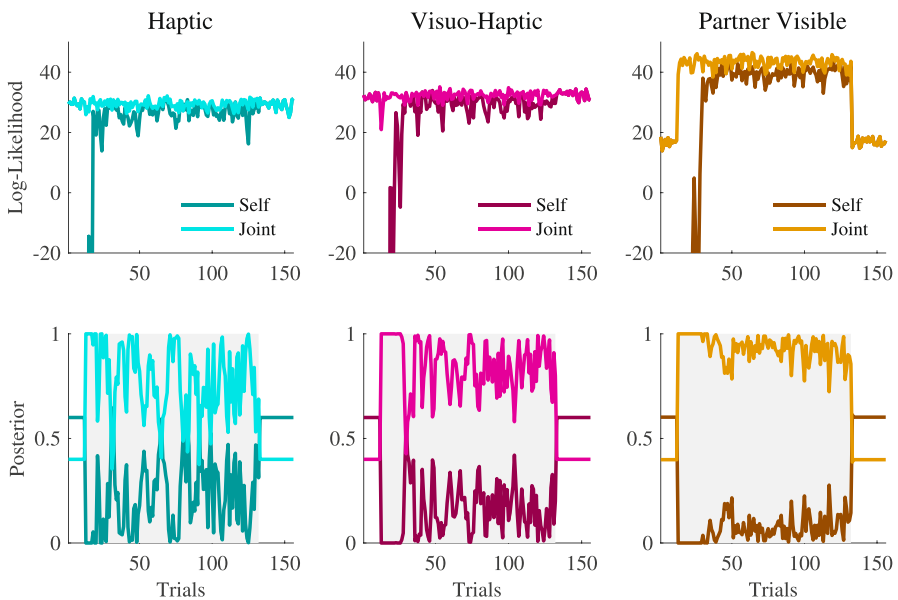
We initially focused on the self and joint-likelihood (and the corresponding self and joint agency beliefs) calculated at the end of each simulated movement (retrospective agency).

The simulation results indicate that during the baseline phase, when the two players are disconnected, the partner model does not contribute significantly to state estimation, hence the self and joint-likelihood are the same. Consequently, the posterior probabilities and thus the retrospective senses of self and joint agency are dominated by the task requirements, encoded by the agency priors. Therefore, retrospective self agency prevails over joint agency.

During the training phase, when the two players are physically connected, the partner model significantly improves the prediction of the sensory outcomes and therefore joint-likelihood is always greater than the self-likelihood. Consequently, the joint posterior probability is greater than its self counterpart, and the joint control action is weighed much more.

In conclusion, interaction with the partner is quickly incorporated in the joint observer and - through fictitious play - in the joint controller. In simulations, we assumed that both players already have a ‘connected’ forward model, i.e. they know already about the connected dynamics. This may be unrealistic, hence in actual experiments the convergence may not be that fast.

Both self and joint-likelihood are affected by uncertainty about the partner. When more reliable information is available to a player – e.g. from the H to the PV



**Fig. 5** Between-trials evolution of self- and joint-likelihoods (top) and self- and joint posterior probabilities (bottom), with increasing amounts (left to right) of available sensory information: haptic (H, left), visuo-haptic (VH, middle) and partner visible (PV, right)

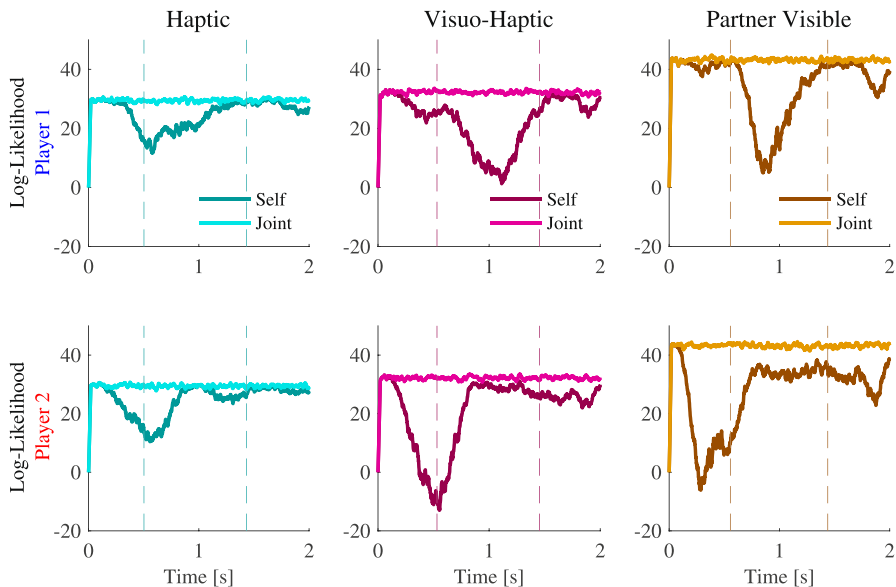
condition – both joint and self-likelihood increase. The prediction of partner actions also becomes more reliable; therefore, the joint-likelihood tends to overcome the self-likelihood. As a consequence, the players increasingly perceive themselves as acting jointly; see Fig. 5.

In conclusion, simulations demonstrate two key model predictions. First, the relative strengths of the sense of self and joint agency are determined by a combination of the task requirements and the match between predicted and observed sensory information. Second, when the information is more reliable the partner model plays a more important role in predicting the sensory outcomes, which strengthens the sense of joint agency with respect to self agency.

### 3.3.2 Self and joint agency within a trial

The proposed model also predicts that both the senses of self and joint agency are time-varying within a single movement. In simulations, we additionally looked at this temporal dynamic. In all three experimental groups we took the last trial within the training phase, when the players have developed a stable joint strategy.

These simulation results are summarized in Fig. 6. The instantaneous joint-likelihood varies little throughout the trial. As expected, it increases when more reliable sensory information is available. This confirms the prediction that reliable sensory information is a major determinant of joint agency.



**Fig. 6** Self- and joint-likelihood of both players within a trial (last trial in the training phase). From left to right: haptic (H), visuo-haptic (VH) and partner visible (PV) groups, for, respectively, Player 1 (top) and Player 2 (bottom). The vertical dashed lines represent the times at which each player crossed their via-point

In contrast, self-likelihood is always smaller and changes within the movement. The discrepancy between self and joint-likelihood reflects the importance of the partner action in estimating the dyad state in different portions of the movement. A large discrepancy implies a stronger partner contribution.

In the simulations, the self-likelihood of both players exhibits a minimum around 300 ms before the time the partner is crossing their own via-point. Interestingly, these results are consistent with previously reported experimental findings. In particular, Chackochan and Sanguinetti (2019) observed that when crossing their own via-point, each player tends to pull their partner – thus behaving like a leader, whereas their partner acts more like a follower.

In our simulations, when the amount of information is greater (e.g. the PV condition), there is a greater gap between self and joint-likelihood, which points at a we-type joint agency (the players tend to lose their sense of self agency as joint agency gets stronger). In contrast, when less information is available (e.g. the H condition), both self- and joint-likelihood remain large throughout the whole trial, thus suggesting a joint agency of shared-type. In conclusion, our simulations predict that when more information about the partner is available, the joint strategy shifts from shared-type to we-type joint agency.

## 4 Discussion

We formulated a comprehensive computational framework to account for the sense of agency in joint action. By extending previous accounts (Fletcher & Frith, 2009; Izawa et al., 2016) of self agency, we suggest that both the senses of self and joint agency can be interpreted in probabilistic terms.

In particular, the prospective (Chambon et al., 2014) and retrospective (Blakemore et al., 2002; Frith et al., 2000; Haggard, 2005) components of the sense of agency can be seen as prior and posterior subjective probabilities (beliefs) of being in control.

Using simple arguments from the optimization framework, widely used in sensorimotor control (Todorov & Jordan, 2002; Daniel M. Wolpert et al., 1995), we first introduce a more general formulation of the comparator model (S. Blakemore et al., 2000; C. D. Frith et al., 2000) for the sense of self agency.

We then argue that in order to address joint action scenarios, the state observer must explicitly account for the partner's ongoing activity (partner model) (Chackochan & Sanguinetti, 2019; Li et al., 2019; Pesquita et al., 2018; Takagi et al., 2017). The partner model is essential for the development of a joint action and a sense of joint agency. It may take different forms and may have different degrees of reliability. Both aspects are central to the development of the sense of joint agency.

We propose differential game theory and non-cooperative games as a general modeling framework to study joint action scenarios which develop in space and time (e.g. involving movements) and are characterized by incomplete information about the partner (Braun et al., 2009; Chackochan & Sanguinetti, 2019; Jarrassé et al., 2012; Li et al., 2019). We propose fictitious play as a general learning mechanism to

describe the gradual development of a stable coordination (Chackochan & Sanguinetti, 2019; Grau-Moya et al., 2013) and the sense of joint agency.

The proposed model is not just a mere mechanistic solution to a control problem. Its level of description is computational (what has to be computed and why) and relies on the well-established notion that the nervous system compares a variety of hypotheses in order to understand the external environment and to take decisions (Gallivan et al., 2017; Heald et al., 2018). Consistent with reports (Bolt & Loehr, 2017; Dewey et al., 2014; Le Bars et al., 2020; Pacherie, 2012) that we maintain distinct senses of self and joint agency, we posited separate observer-controller pairs, one assuming that we are acting alone ('self' loop) and another assuming that we are acting jointly ('joint' loop). The loops run continuously and in parallel. The two observers provide continuously updated subjective assessments of the 'self' (*I am acting alone*) and 'joint' (*I am acting jointly with someone else*) hypotheses, which in turn determine action selection. Each observer provides a specific interpretation. The 'agency' signal (posterior probability of either self or joint) is used to decide which controller is appropriate for the current scenario (Haruno et al., 2001). The model can be easily extended to address more scenarios, in which we just observe someone else's action ('other' scenario) or interaction with multiple agents, where the partner model combines the contributions of multiple partners (Takagi et al., 2019).

## 4.1 Model predictions

To clarify how the model works and to explore less intuitive model predictions and implications, we simulated a joint action experiment from a published study (Chackochan & Sanguinetti, 2019).

### 4.1.1 Self and joint agency are affected by the quality of sensory information

The extended probabilistic formulation of the comparator model (Frith., 2005; C. D. Frith et al., 2000; Pacherie, 2008) predicts that the amount of information available about the dyad (player and partner) modulates our sense of self and joint agency. This prediction is demonstrated by the model simulations: when sensory uncertainty is reduced, the players' ability to predict their partners' actions improves, and they feel more strongly to be part of a group, thus leading to a stronger sense of joint agency. This is consistent with reports that humans feel a stronger sense of joint agency when they coordinate their actions with a more predictable partner (Bolt & Loehr, 2017).

### 4.1.2 Are there different types of joint agency?

The proposed model posits that we experience distinct senses of self and joint agency, but they are not independent. Pacherie (2012) argued that there are two qualitatively different types of sense of joint agency. In 'we-agency' (Gallotti & Frith, 2013; Pacherie, 2012), strong sense of joint agency is associated to a weak

sense of self agency. In ‘shared-agency’ a strong sense of joint agency is compatible with a strong sense of self agency (Loehr, 2018; Pacherie, 2012).

Our proposed model captures both scenarios, with no need to assume different ‘types’ or ‘modes’ of the sense of joint agency (Pacherie, 2012). We suggest that the relation between the senses of joint and self agency is determined by the task requirements and the predictability of partner’s actions, with no need to assume qualitatively different ‘modes’ of the sense of joint agency. In particular, the ‘we-agency’ scenarios corresponds to situations – or phases of a joint action – in which our partner’s actions have a strong impact on our sensory afferences. In this case, the joint-likelihood would be much greater than the self-likelihood, which would place a bias toward joint agency. For instance, singers in a choir with different voices coordinate their actions to achieve a shared goal and equally contribute in its fulfillment. As such they are likely to experience a greater sense of joint agency and a comparatively lower sense of self agency as each individual only contributes a fraction of the overall performance. Conversely, ‘shared-agency’ scenarios correspond to situations in which partner actions play a weaker role in the overall coordination, or have different impacts in different phases of the movement, e.g. when the players have different roles. In this case the partner information would not add much to the joint-likelihood with respect to the self-likelihood.

Similarly, the model predicts that we experience different proportions of sense of joint and self agency in competitive situations with respect to collaborative scenarios. For instance, two fighters likely experience a lower sense of joint agency with respect to two partners lifting a heavy object together. Joint weight lifting implies a strong prospective sense of joint agency and a high joint-likelihood, hence a strong retrospective sense of joint agency as compared to self-agency. Fighting implies a strong prospective sense of joint agency, but the joint-likelihood would be low, as the competitor would aim at being less predictable. As a consequence, fighting implies a weaker retrospective sense of joint agency (as compared to self-agency). These predictions are also consistent with (Silver et al., 2021), which identify the level of cooperation as the key action feature which modulates the sense of joint agency.

The model provides a computational substrate for the concepts of we-ness and we-representation (Gallotti & Frith, 2013; Kourtis et al., 2019). A partner in a joint action gathers significance only if he/she is seen in the self-perspective, as an augmentation of the self. We would not need to account for our partner if there are no requirements to undergo a joint action, as in the unconnected phases of the simulated experiment.

#### 4.1.3 Agency as a dynamic experience

Our proposed model suggests that both senses of joint and self agency evolve in time. Before an action takes place, an early (prospective) sense of agency is determined by the task goal, requirements and constraints, and additional cognitive and emotional factors, as suggested by Synofzik et al. (2013) and Chambon et al. (2014).

This initial bias – the self and joint priors in our model – is combined with self and joint-likelihood signals, each quantifying the reliability of a different predictor of the ongoing action. The relative magnitudes of self and joint-likelihood change within a single action, up to action completion (retrospective agency). In the simulations, each player coordinates with their partner, but also needs to maintain their identity to achieve their own sub-goal – i.e. crossing their respective via-point. When the sensory information is more reliable, the sense of joint agency becomes stronger. However, the sense of self agency only increases if the partner is not perceived to effectively contribute to the action.

#### 4.1.4 Agency determines the decision/control process

In our proposed multi-loop control architecture, the experienced self and joint agency play a pivotal role in motor command selection and therefore in the development of the joint strategy. In other words, the sense of agency is not just about perception, but also affects action. The ‘self’ and ‘joint’ state observers give rise to separate controllers, whose relative importance is weighed by the sense of self and joint agency. Hence the sense of agency (self and joint) is not simply a byproduct of perception, but emerges from this dynamic process which involves a bidirectional interplay between action control and estimation of own and partner state.

A strong sense of joint agency places more emphasis on more collaborative actions and facilitates the convergence to a stable coordination, in which each player accounts for their partner and selects their action accordingly. If information about the partner is more uncertain, the sense of joint agency is lower and the players converge to a strategy which relies less on partner’s actions. In both cases, the selected actions further emphasize either the senses of joint or self agency.

## 4.2 Learning in joint action

An open issue in joint action is how two players converge to a coordination strategy. Fictitious play assumes that the players select the action which maximizes their respective objective function, by accounting for the empirical probability distributions of the actions of their opponent. A key property of fictitious play is that it only requires that each agent knows their partner’s probability distribution of previous actions. Higher-order theories of mind, possibly representing partner intentions (i.e. the control policies) or goals (i.e. their objective function) would lead to more efficient learning and/or to different Nash equilibria. Higher-order forms of theory of mind may lead to an infinite regression - each agent represents their opponent’s policy which accounts for their own and so on; see also Yoshida et al. (2008).

If an agent does not know in advance the relation between action and its subjective ‘value’ (in terms of our model, the ‘task representation’ block), he/she can learn it through some biologically plausible form of reinforcement learning, like Q-learning (Claus & Boutilier, 1998). In joint action scenarios, this is more complicated as

each agent's goal – and its subjective value - depends on the actions of both agents. Here a question arises of whether learning the action(s)-value and predicting the partner actions occur independently, or are actually part of one single process. In a recent study focusing on sensorimotor versions of classical discrete games, Lindig-León et al. (2021) observed that convergence to a Nash equilibrium is consistent with a model-free form of reinforcement learning, in which actions are generated as a trade-off between their value and the requirement of minimizing their change with respect to the previous trial. This learning mechanism does not explicitly account for partner actions, but it is unclear if it would extend to more complex forms of coordination that involve more than just discrete decisions.

In conclusion, how a coordination strategy is learned is an open issue, and the value of the fictitious play hypothesis is that it represents a reference hypothesis against which other possibilities can be tested empirically.

### 4.3 Relation with previous models of joint action

As noted in section 2.2.3, the proposed model is inspired to the MOSAIC computational architecture (D. M. Wolpert & Kawato, 1998), which was originally developed to address motor system modularity. MOSAIC posits multiple observer-controller pairs, each specialized for specific actions (e.g. grasping a specific object) or environmental conditions. In addition to the observer and the controller, each module has a responsibility predictor which uses contextual information to estimate the prior of that module. Priors and likelihoods (provided by the observer) are then combined into a weighing signal for the motor commands calculated by each controller module. The model was later extended to address sensorimotor learning (Haruno et al., 2001) and action observation (Wolpert et al., 2003). This latter study suggested that one pair of such models – one for action control, and the other devoted to action observation - could capture the mechanisms underlying social interaction. A recent study (Haar & Donchin, 2019) discusses MOSAIC's possible neural substrates.

Here we extend the MOSAIC model by (i) incorporating priors which, combined with the likelihoods, provide the posterior probability of each scenario; and (ii) allowing for loops reflecting self and joint action. Self and joint action are treated as different scenarios, and self and joint agency are determined as their respective posterior probabilities. Similar to 'responsibilities' in the MOSAIC model, the agency signals affect the action selection process.

The Predictive Joint Action Model, PJAM (Pesquita et al., 2018) specifically addresses joint action. PJAM builds on HMOSAIC, an evolution of the original model (Haruno et al., 2003) which addresses the multilevel character of perception and action through a hierarchy of MOSAIC layers. Each layer receives posterior probabilities from the lower level module (bottom-up path), which specify the currently selected module in the current behavioral situation. The output of higher-level modules is a set of (top-down) prior probabilities of the subordinate modules, which act to prioritize lower-level module selection.

PJAM posits a hierarchy of 'predictive processors' (i.e. observer-controller pairs), which captures multiple levels of action representation, from high-level (symbolic)

to low-level (movement feature). Each layer posits a ‘self’ and a ‘other’ observer-controller loop. At higher levels in the hierarchy the loops merge into a ‘joint’ loop, whereas at lower levels of the hierarchy the ‘self’ and ‘other’ loops remain distinct. Although the architecture is only described qualitatively, it seems effective in capturing situations in which the incentive to act together is encoded into each player’s objective function and is evaluated before the action takes place. PJAM is not intended for generating motor commands in a joint action context, but rather as a module for hierarchical predictive processing, focusing on action planning.

With respect to PJAM, our proposed model focuses on the generation of motor commands aimed at gradually achieving a stable coordination with a partner, even in situations in which there is minimal or no information about partner’s actions or intentions. This necessarily requires a ‘joint’ controller, which explicitly accounts for both the ‘self’ and ‘other’ states. We also suggest that the controller block is an optimal feedback controller, a more general formulation than MOSAIC or PJAM, which is easily extended through differential game theory to address joint action controllers.

Likewise, on the prediction side we point out that in general the ‘self’ and ‘other’ state predictions cannot run independently, even at lower levels of representation – think for instance of mechanical coupling. This suggests that a general hierarchical predictor-observer architecture must involve a ‘joint’ loop at all levels of representation. In other words, a ‘joint’ predictive processor (or equivalently two inter-dependent self-other predictive processors) is required at all levels of the hierarchy.

In our model, we added a ‘self’ observer-controller loop to the ‘joint’ loop to account for the empirical observation of distinct experiences of joint and self agency. In conclusion, our proposed model differs from PJAM in that it posits a ‘joint’ and ‘self’ rather than a ‘self’ and ‘other’ loop. Another difference is that our model does not address the multilevel character of action representation and only focuses on the lower level (movement feature), but the extension is straightforward.

Finally, and distinctively, we suggest that the responsibility signals can be interpreted as distinct subjective ‘joint’ and ‘self’ agency beliefs, an aspect which is not specifically addressed by HMOSAIC or PJAM.

In conclusion, our proposed model departs significantly from PJAM. We believe it provides a coherent picture of the inter-relation of prediction, control, and the sense of agency in a broader range of joint actions.

Recently, in the context of a gaze-contingent virtual task, Brandi et al. (2019) proposed a Bayesian model for “social agency”, defined as the experience of control over the social environment. Based on experimental work of Pfeiffer et al. (2012), the model relies on specific key features of partner action, namely gaze direction (e.g. making eye contact) and temporal responsiveness. They propose that the (precision-weighted) mismatch between these observations and their predictions determines the degree of social agency. The precision-weighted sensory prediction error can be interpreted as the logarithm of the likelihood, hence the model is very related to the one proposed here. In fact, in a Bayesian framework the specific features that are more relevant to joint agency automatically emerge as the main contributors to the likelihood, with no need to specify them a priori.



Our proposed model can be related to the free energy principle (FEP) framework (Friston et al., 2006). FEP aims at unifying perception and action in terms of the minimization of one single quantity – free energy. Like the FEP, our proposed model relies on an optimization framework, but assumes separate costs for action generation and perception. Recently, Kahl and Kopp (2018) proposed a hierarchical FEP-based model of sensorimotor coordination, which is capable to discriminate between self and other actions. Self and other agency are simulated in a handwriting task, by manipulating perception so that it is either consistent or not with their own action. They found that inconsistent (i.e. other) behavior corresponds to a greater free energy magnitude. However, the model does not directly address joint action and it is unclear how it would select the correct action in ‘joint’ scenarios in which the players are required to act at the same time, like in continuous coordination games.

All the above models (Brandi et al., 2019; Haruno et al., 2003; Kahl & Kopp, 2018; Pesquita et al., 2018) propose a hierarchical architecture to address the different levels of abstraction of the perception-action process. Here we only focus on a minimal architecture which focuses on the bottom layer of the prediction-action cycle, but the proposed framework can be extended to multiple layers to account for the integration of partner models at different levels of description.

## 5 Conclusions

Based on a Bayesian probabilistic framework we propose a general model of joint action which brings together prediction, control and the sense of agency in a broad range of joint action scenarios. The model extends previous models of the sense of self agency with two main extensions: a partner model which is capable of predict the partner actions or intentions as part of a ‘joint’ state observer, and an optimal controller based on differential game theory. The model posits two observer-controller loops which run in parallel, accounting for ‘joint’ and ‘self’ prediction and control and provide continuous estimates of the experienced sense of joint and self agency. The proposed probabilistic framework captures the inter-individual variability in joint action, so that the sense of agency is a highly subjective experience. Variability affects the sensorimotor loop at different levels. Individual differences are apparent in sensory acuity and in the accuracy of our internal representations, and personal traits affect the subjective evaluation of goals and task (Vallacher & Wegner, 1989), leading to a more distal or proximal experience of action. These features are accounted by the priors and by the reliability of the sensory channels.

The proposed model has several speculative aspects which call for empirical test. The nature of the partner (‘other’) model is debated. It may only focus on ongoing movements; or the ongoing motor commands; or even the underlying intentions/goals. This aspect probably depends on the task and its context, and can be explored empirically. Another aspect is the relation of the partner model with the prediction of our own action (‘self’ observer). Different to similar accounts (Pesquita et al., 2018) we suggest that ‘self’ and ‘other’ observer are interdependent, so that they should be really considered as a single ‘joint’ observer.

How is a stable coordination strategy developed is another aspect of the proposed model. We proposed fictitious play as a general mechanism, but there are alternatives – e.g. reinforcement learning – which may better capture the experimental observations and could provide additional insights about the prospective and retrospective components of the sense of agency.

Another aspect of the model that call for additional empirical test is the introduction of a ‘self’ observer-controller loop as the basis to account for the empirical observations of distinct senses of joint and self agency: how does it compare with the accounts of we-agency and shared-agency subtypes of joint agency?

Differential game theory in conjunction with optimal estimations provides a wide range of tools which can be used to address joint action.

The design of experiments to disentangle the senses of self and joint agency is extremely challenging. Questionnaires provide useful tools to quantify the sense of self and joint agency, but they have limitations in capturing the temporal evolution of the agency during a sustained joint action performance and in the independent assessment of joint and self agency. Other approaches, like intentional binding (Haggard et al., 2002), need to be extended to specifically address joint agency.

Overall, computational and empirical approaches can potentially inform each other for a complete understanding and a general description of the sense of agency and its behavioral implications.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11097-022-09819-5>.

**Code availability** Code for model simulation will be made available upon reasonable request.

**Author’s contributions** Conceptualization: Cecilia De Vicariis, Vinil T. Chackochan, Vittorio Sanguineti.

Data curation: Cecilia De Vicariis.

Formal analysis: Vittorio Sanguineti, Cecilia De Vicariis.

Software: Cecilia De Vicariis, Vittorio Sanguineti.

Supervision: Vittorio Sanguineti.

Validation: Vittorio Sanguineti.

Visualization: Cecilia De Vicariis, Vinil T. Chackochan.

Writing – original draft: Cecilia De Vicariis.

Writing – review & editing: Cecilia De Vicariis, Vinil T. Chackochan, Vittorio Sanguineti.

**Funding** Open access funding provided by Università degli Studi di Genova within the CRUI-CARE Agreement. The work is partly supported by a grant (ModuLimb) from the Italian Ministry of Education, University and Research (MIUR) – Research Project of National Interest (PRIN).

**Data availability** Not applicable.

## Declarations

**Conflict of interest** No

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the

material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bacharach, M. (1999). Interactive team reasoning: A contribution to the theory of co-operation. *Research in Economics*, 53(2), 117–147. <https://doi.org/10.1006/reec.1999.0188>
- Başar, T., & Olsder, G. J. (1999). *Dynamic noncooperative game theory* (Vol. 23, 2nd ed.p. SIAM).
- Berger, U. (2007). Brown's original fictitious play. *Journal of Economic Theory*, 135(1), 572–578. <https://doi.org/10.1016/j.jet.2005.12.010>
- Blakemore, S., Wolpert, C. A. D., & Frith, C. (2000). Why can't you tickle yourself?, 10–16. <https://doi.org/10.1097/00001756-200008030-00002>.
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, 6(6), 237–242. [https://doi.org/10.1007/978-3-319-57111-9\\_1389](https://doi.org/10.1007/978-3-319-57111-9_1389)
- Bolt, N. K., & Loehr, J. D. (2017). The predictability of a partner's actions modulates the sense of joint agency. *Cognition*, 161, 60–65. <https://doi.org/10.1016/j.cognition.2017.01.004>
- Bolt, N. K., Poncelet, E. M., Schultz, B. G., & Loehr, J. D. (2016). Mutual coordination strengthens the sense of joint agency in cooperative joint action. *Consciousness and Cognition*, 46, 173–187. <https://doi.org/10.1016/j.concog.2016.10.001>
- Brandi, M. L., Kaifel, D., Bolis, D., & Schilbach, L. (2019). The interactive self - a review on simulating social interactions to understand the mechanisms of social agency. *I-Com*, 18(1), 17–31. <https://doi.org/10.1515/icom-2018-0018>
- Braun, D. A., Ortega, P. A., & Wolpert, D. M. (2009). Nash equilibria in multi-agent motor interactions. *PLoS Computational Biology*, 5(8). <https://doi.org/10.1371/journal.pcbi.1000468>
- Brown, G. W. (1951). Iterative solution of games by fictitious play. *Activity Analysis of Production and Allocation*. Wiley.
- Carruthers, P., & Smith, P. K. (1996). Theories of theories of mind. In: (Eds.). Cambridge University Press.
- Chackochan, V. T., & Sanguineti, V. (2019). Incomplete information about the partner affects the development of collaborative strategies in joint action. *PLoS Computational Biology*, 15(12), e1006385. <https://doi.org/10.1371/journal.pcbi.1006385>
- Chambon, Sidarus, & Haggard. (2014). From action intentions to action effects: How does the sense of agency come about? *Frontiers in Human Neuroscience*, 8. <https://doi.org/10.3389/fnhum.2014.00320>
- Claus, C., & Boutillier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. *AAAI/IAAI*, 1998(746–752), 2.
- Dewey, J. A., Pacherie, E., & Knoblich, G. (2014). The phenomenology of controlling a moving object with another person. *Cognition*, 132(3), 383–397. <https://doi.org/10.1016/j.cognition.2014.05.002>
- Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, 10(1), 48–58. <https://doi.org/10.1038/nrn2536>
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris*, 100(1–3), 70–87. <https://doi.org/10.1016/j.jphysparis.2006.10.001>
- Frith, C. D., & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, 60(3), 503–510. <https://doi.org/10.1016/j.neuron.2008.10.032>
- Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 355(1404), 1771–1788. <https://doi.org/10.1098/rstb.2000.0734>
- Frith. (2005). The self in action: Lessons from delusions of control. *Consciousness and Cognition*, 14(4), 752–770. <https://doi.org/10.1016/j.concog.2005.04.002>
- Fudenberg, D., & Levine, D. (1998). Learning in games. *European Economic Review*, 42(3–5), 631–639. [https://doi.org/10.1016/S0014-2921\(98\)00011-7](https://doi.org/10.1016/S0014-2921(98)00011-7)
- Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17(2), 535–543. <https://doi.org/10.1016/j.concog.2008.03.003>

- Gallese, V., & Goldman, A. (1998). Mirror neurons and the theory of mind rreading. *Trends in Cognitive Sciences*, 2(12), 493–501.
- Gallivan, J. P., Stewart, B. M., Baugh, L. A., Wolpert, D. M., & Flanagan, J. R. (2017). Rapid automatic motor encoding of competing reach options. *Cell Reports*, 18(7), 1619–1626. <https://doi.org/10.1016/j.celrep.2017.01.049>
- Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, 17(4), 160–165. <https://doi.org/10.1016/j.tics.2013.02.002>
- Ganesh, G., Takagi, A., Osu, R., Yoshioka, T., Kawato, M., & Burdet, E. (2014). Two is better than one: Physical interactions improve motor performance in humans. *Scientific Reports*, 4(1), 1–7. <https://doi.org/10.1038/srep03824>
- Gangopadhyay, N., & Miyahara, K. (2015). Perception and the problem of access to other minds. *Philosophical Psychology*, 28(5), 695–714. <https://doi.org/10.1080/09515089.2014.895935>
- Gillijns, S., & De Moor, B. (2007). Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, 43(1), 111–116. <https://doi.org/10.1016/j.automatica.2006.08.002>
- Grau-Moya, J., Hez, E., Pezzulo, G., & Braun, D. A. (2013). The effect of model uncertainty on cooperation in sensorimotor interactions. *Journal of the Royal Society Interface*, 10(87). <https://doi.org/10.1098/rsif.2013.0554>
- Haar, S., & Donchin, O. (2019). A revised computational neuroanatomy for motor control. *Journal of Cognitive Neuroscience*. <https://doi.org/10.1007/s00221-008-1280-5>
- Haggard, P. (2005). Conscious intention and motor cognition. 9(6). <https://doi.org/10.1016/j.tics.2005.04.012>.
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5(4), 382–385. <https://doi.org/10.1038/mn827>
- Harsanyi, J. C. (1967). Games with incomplete information played by “Bayesian” players. *Management Science*, 14(12 SUPPL), 1804–1824. <https://doi.org/10.1287/mnsc.1040.0270>
- Haruno, M., Wolpert, D. M., & Kawato, M. (2001). MOSAIC model for sensorimotor learning and control. *Neural Computation*, 13(10), 2201–2220. <https://doi.org/10.1162/089976601750541778>
- Haruno, M., Wolpert, D. M., & Kawato, M. (2003). Hierarchical MOSAIC for movement generation. *International Congress Series*, 1250(C), 575–590. [https://doi.org/10.1016/S0531-5131\(03\)00190-0](https://doi.org/10.1016/S0531-5131(03)00190-0)
- Heald, J. B., Ingram, J. N., Flanagan, J. R., & Wolpert, D. M. (2018). Multiple motor memories are learned to control different points on a tool. *Nature Human Behaviour*, 2(4), 300–311. <https://doi.org/10.1038/s41562-018-0324-5>
- Izawa, J., Asai, T., & Imamizu, H. (2016). Computational motor control as a window to understanding schizophrenia. *Neuroscience Research*, 104, 44–51. <https://doi.org/10.1016/j.neures.2015.11.004>
- Jarrassé, N., Charalambous, T., & Burdet, E. (2012). A framework to describe, analyze and generate interactive motor behaviors. *PLoS One*, 7(11). <https://doi.org/10.1371/journal.pone.0049945>
- Kahl, S., & Kopp, S. (2018). A predictive processing model of perception and action for self-other distinction. *Frontiers in Psychology*, 9, 1–6. <https://doi.org/10.3389/fpsyg.2018.02421>
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45. <https://doi.org/10.1115/1.3662552>
- Knoblich, G., Butterfill, S., & Sebanz, N. (2011). *Psychological research on joint action : Theory and data* (Vol. 54). <https://doi.org/10.1016/B978-0-12-385527-5.00003-6>
- Kourtis, D., Woźniak, M., Sebanz, N., & Knoblich, G. (2019). Evidence for we-representations during joint action planning. *Neuropsychologia*, 131, 73–83. <https://doi.org/10.1016/j.neuropsychologia.2019.05.029>
- Lafleur, A., Soulières, I., & Forgeot d’Arc, B. (2020). Sense of agency: Sensorimotor signals and social context are differentially weighed at implicit and explicit levels. *Consciousness and Cognition*, 84. <https://doi.org/10.1016/j.concog.2020.103004>
- Le Bars, S., Devaux, A., Nevidal, T., Chambon, V., & Pacherie, E. (2020). Agents’ pivotality and reward fairness modulate sense of agency in cooperative joint action. *Cognition*, 195, 104117. <https://doi.org/10.1016/j.cognition.2019.104117>
- Li, Y., Carboni, G., Gonzalez, F., Campolo, D., & Burdet, E. (2019). Differential game theory for versatile physical human–robot interaction. *Nature Machine Intelligence*, 1(1), 36–43. <https://doi.org/10.1038/s42256-018-0010-3>
- Lindig-León, C., Schmid, G., & Braun, D. A. (2021). Nash equilibria in human sensorimotor interactions explained by Q-learning with intrinsic costs. *Scientific Reports*, 11(1), 1–15. <https://doi.org/10.1038/s41598-021-99428-0>

- Loehr, J. D. (2018). Shared credit for shared success: Successful joint performance strengthens the sense of joint agency. *Consciousness and Cognition*, 66, 79–90. <https://doi.org/10.1016/j.concog.2018.11.001>
- Meinhardt-Injac, B., Daum, M. M., Meinhardt, G., & Persike, M. (2018). The two-systems account of theory of mind: Testing the links to social-perceptual and cognitive abilities. *Frontiers in Human Neuroscience*, 12, 1–12. <https://doi.org/10.3389/fnhum.2018.00025>
- Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, 21(1), 59–68. <https://doi.org/10.1016/j.concog.2011.08.010>
- Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, 17(1), 136–144. <https://doi.org/10.1016/j.concog.2006.12.004>
- Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, 54(2), 286–295. <https://doi.org/10.2307/1969529>
- Nash, J. (1953). Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, 128–140.
- Newton, J. (2018). Evolutionary game theory: A renaissance. *Games*, 9(2). <https://doi.org/10.3390/g9020031>
- Noy, L., Weiser, N., & Friedman, J. (2017). Synchrony in joint action is directed by each participant's motor control system. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00531>
- Pacherie, E. (2008). The phenomenology of action : A conceptual framework. *Cognition, Elsevier*, 107(1), 179–217.
- Pacherie, E. (2012). The Phenomenology of Joint Action: Self-Agency vs Joint Agency. *Joint Attention: New Developments*. MIT Press. <https://doi.org/10.7551/mitpress/8841.003.0017>.
- Pacherie, E. (2014). How does it feel to act together? *Phenomenology and the Cognitive Sciences*, 13(1), 25–46. <https://doi.org/10.1007/s11097-013-9329-8>
- Pesquita, A., Whittlewell, R. L., & Enns, J. T. (2018). Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychonomic Bulletin and Review*, 25(5), 1751–1769. <https://doi.org/10.3758/s13423-017-1393-6>
- Pfeiffer, U., Schilbach, L., Timmermans, B., Jording, M., Bente, G., & Vogeley, K. (2012). Eyes on the mind: Investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in Psychology*, 3, 537.
- Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6), 1234–1246. <https://doi.org/10.1037/0096-1523.31.6.1234>
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10(2), 70–76. <https://doi.org/10.1016/j.tics.2005.12.009>
- Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, 185(3), 359–381. <https://doi.org/10.1007/s00221-008-1280-5>
- Silver, C. A., Tatler, B. W., Chakravarthi, R., & Timmermans, B. (2021). Social agency as a continuum. *Psychonomic Bulletin and Review*, 28(2), 434–453. <https://doi.org/10.3758/s13423-020-01845-1>
- Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, 17(1), 219–239. <https://doi.org/10.1016/j.concog.2007.03.010>
- Synofzik, Vosgerau, & Voss. (2013). The experience of agency: An interplay between prediction and postdiction. *Frontiers in Psychology*, 4, 1–8. <https://doi.org/10.3389/fpsyg.2013.00127>
- Takagi, A., Ganesh, G., Yoshioka, T. et al. (2017). Physically interacting individuals estimate the partner's goal to enhance their movements. *Nat Hum Behav*, 1, 0054. <https://doi.org/10.1038/s41562-017-0054>
- Takagi, A., Usai, F., Ganesh, G., Sanguineti, V., & Burdet, E. (2018). Haptic communication between humans is tuned by the hard or soft mechanics of interaction. *PLoS Computational Biology*, 14(3), 1–17. <https://doi.org/10.1371/journal.pcbi.1005971>
- Takagi, A., Hirashima, M., Nozaki, D., & Burdet, E. (2019). Individuals physically interacting in a group rapidly coordinate their movement by estimating the collective goal. *ELife*, 8, 1–19. <https://doi.org/10.7554/eLife.41328>
- Todorov, E. (2004). Optimality principle in sensorimotor control. *Nature Neuroscience*, 7(9), 907–915.
- Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11), 1226–1235. <https://doi.org/10.1038/nm963>
- Tuomela, R. (2007). *The philosophy of sociality: The shared point of view*. Oxford University Press.
- Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, 57(4), 660–671. <https://doi.org/10.1037/0022-3514.57.4.660>

- van der Wel, R. P. R. D. (2015). Me and we: Metacognition and performance evaluation of joint actions. *Cognition*, *140*, 49–59. <https://doi.org/10.1016/j.cognition.2015.03.011>
- Vesper, C., Abramova, E., Büttepage, J., Ciardo, F., Crossey, B., Effenberg, A., et al. (2017). Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in Psychology*, *7*. <https://doi.org/10.3389/fpsyg.2016.02039>
- Wegner, D. M. (2003). The mind's best trick: How we experience conscious will. *Trends in Cognitive Sciences*, *7*(2), 65–69. [https://doi.org/10.1016/S1364-6613\(03\)00002-0](https://doi.org/10.1016/S1364-6613(03)00002-0)
- Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*(7), 480–492. <https://doi.org/10.1037/0003-066X.54.7.480>
- Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*(6), 209–216. <https://doi.org/10.1016/B978-008045046-9.01311-5>
- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*(7–8), 1317–1329. [https://doi.org/10.1016/S0893-6080\(98\)00066-5](https://doi.org/10.1016/S0893-6080(98)00066-5)
- Wolpert, D. M., & Miall, R. C. (1996). Forward models for physiological motor control. *Neural Networks*, *9*(8), 1265–1279.
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882. <https://doi.org/10.1126/science.7569931>
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B*, 593–602. <https://doi.org/10.1098/rstb.2002.1238>
- Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. *PLoS Computational Biology*, *4*(12). <https://doi.org/10.1371/journal.pcbi.1000254>
- Zalla, T., & Sperduti, M. (2015). The sense of agency in autism spectrum disorders: A dissociation between prospective and retrospective mechanisms? *Frontiers in Psychology*, *6*, 1–7. <https://doi.org/10.3389/fpsyg.2015.01278>
- Zunino, A., Cavazza, J., Volpi, R., Morerio, P., Cavallo, A., Becchio, C., & Murino, V. (2020). Predicting intentions from motion: The subject-adversarial adaptation approach. *International Journal of Computer Vision*, *128*(1), 220–239. <https://doi.org/10.1007/s11263-019-01234-9>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Cecilia De Vicariis<sup>1</sup>  · Vinil T. Chackochan<sup>2,3</sup> · Vittorio Sanguineti<sup>1</sup>

✉ Cecilia De Vicariis  
ceciliadevicariis@gmail.com

<sup>1</sup> Department of Informatics, Bioengineering, Robotics and System Engineering Genoa (ITALY), University of Genoa, Genoa, Italy

<sup>2</sup> Bournemouth University, Bournemouth, UK

<sup>3</sup> Salisbury District Hospital, Salisbury, UK