# SLOW BUT FLEXIBLE OR FAST BUT RIGID? DISCRETE AND CONTINUOUS PROCESSES COMPARED

**Matteo Priorelli**
Institute of Cognitive Sciences and Technologies
National Research Council of Italy
Padova, Italy

**Ivilin Peev Stoianov**
Institute of Cognitive Sciences and Technologies
National Research Council of Italy
Padova, Italy
ivilinpeev.stoianov@cnr.it

**ABSTRACT**

A tradeoff always exists when dealing with multi-step tasks. High-level cognitive processes can find the best sequence of actions to achieve goals in uncertain environments, but they are slow and require significant computational demand. Contrarily, lower-level processing allows reacting to environmental stimuli rapidly, but with limited capacity to determine optimal actions. Through reiteration of the same task, biological organisms find the optimal tradeoff: from primitive movements, composite actions gradually emerge by creating task-specific neural structures. The two frameworks of a recent theory called "active inference" can capture high-level and low-level processes of human behavior, but how task specialization may occur in there terms is still unclear. Here, we compare two hierarchical strategies on a pick-and-place task: a discrete-continuous model with planning capabilities and a continuous-only model with fixed transitions. We analyze several consequences of defining movements in intrinsic and extrinsic domains. Finally, we propose how discrete actions might be encoded into continuous representations, comparing them with different motor learning phases and laying the foundations for further studies on bio-inspired task adaptation.

## 1 Introduction

Real-world tasks such as picking and placing an object involve complex actions comprising multiple primitive movements. How does the brain support efficient execution of such actions? While a movement sequence is easy to plan in static contexts, a difficulty emerges when acting in dynamic environments, e.g., when an object has to be grasped on the fly. Tackling multi-step tasks is known to occur deep in the cortical hierarchy, by areas processing discrete, slow-varying entities [1]. But the advantage of discrete representations – useful for planning [2] – comes at a cost: if an object moves too fast, deep processing might be too slow for the grasping action to succeed.

Motor skill consolidation is a well-known phenomenon [3], clearly evident in athletes [4, 5, 6]: during repeated exposure to a task, an initial learning phase leaves its place to autonomous movements as the athlete becomes more proficient [7]. Cortical involvement is gradually reduced as specialized neural structures are constructed in lower, fast subcortical regions [8], including fine adjustment in the spinal cord reflexes [9]. These structures cannot extract invariant representations to perform high-level decision-making but they react rapidly to the sensorium. The basal ganglia are known to be central to task specialization as the striatum encodes chunked representations of action steps composing a learned habit [10, 11]. But this change seems to be much more pervasive, involving information processing within the cortex itself and a general activity shift from anterior to posterior regions [12], as evidenced by plastic changes in the primary motor cortex [13]. Nonetheless, the underlying brain mechanisms are still unclear, and understanding their computational basis is compelling.

An interesting proposal is that the brain maintains a model of the task dynamics, and skill consolidation consists of fine-tuning this model to achieve a stable behavior accounting for environmental uncertainty [14, 15]. This hypothesis fits with a recent theory called active inference, which brings insights of increasing appeal into the computational role of the nervous system [16, 17, 18]. Active inference provides a formalization of the two essential components of control, sensing and acting, which supposedly aim to resolve the critical goal of all organisms: to survive in uncertain environments by operating within preferred states (e.g., maintaining a constant temperature). As in predictive coding [19], active inference assumes that organisms perceive the environment through an internal generative model constructed by inferring how hidden causes produce sensations [17, 20]. Perception and action perform a minimization of a quantity

called *free energy*: while perception gradually adjusts internal expectations to match sensory evidence, action gradually samples those sensations that make the expectations true [17, 20].

The two frameworks of this theory have been used to analyze high-level and low-level processes of human behavior under a unified perspective. Active inference in continuous time solves motor tasks consisting of simple trajectories by defining an internal dynamic model about external targets [20, 21]. Elementary movements are then resolved by local suppression of proprioceptive prediction errors [22], forcing a physical change in the actuators. However, when it comes to multi-step movements, the continuous framework falls short since the minimization only deals with current sensory signals.

At the other extreme, decision-making is addressed by the discrete formulation [23, 24, 25], which views planning as an inferential mechanism [26, 27] that leverages Partially Observed Markov Decision Processes (POMDPs) to make predictions about future states. This is possible because discrete models minimize the free energy of expected observations [28, 29, 30, 31], and it seems that cortical operations can be simulated by such models [1] with higher biological plausibility compared to reinforcement learning [32, 33]. The downside of discrete representations is the lack of real-time interaction with the environment.

Between the two stands a third architecture: a hybrid framework, which is a discrete model combined with its continuous counterpart [34, 35]. This involves the transformation of continuous signals into discrete messages through the so-called Bayesian Model Reduction (BMR) [36, 37]. Solving multi-step tasks with a hybrid model implies associating elementary continuous trajectories to discrete states for planning. This approach has not received adequate attention yet, as not many implementations can be found in the literature [2, 18, 34, 38, 35, 39, 40], none simulating applications in dynamic contexts.

Most importantly, it is unclear how active inference could represent the virtuous cycle of motor skill acquisition, which begins with the rigid information processing of subcortical structures and culminates with the emergence of task-specific structures. In this view, the hybrid framework could be a useful starting point: discrete computations can simulate the behavior of an agent that has no prior idea of what actions to take. However, as soon as the agent has learned a policy that correspond to a sequence of a-priori-defined actions, it can adapt by encoding the transitions between discrete states into continuous dynamics, which then represent a learned motor skill.

## 2 Results

To shed light on task specialization under active inference, we propose a way to define and tackle dynamic multi-step tasks, and compare the behavioral differences of discrete and continuous frameworks, showing that discrete actions and continuous trajectories work under similar mechanisms. To this aim, we considered a pick-and-place operation. The agent's body is an 8-DoF simulated arm with two fingers. At each trial, a target object spawns with random and unknown position, and the agent has to reach the object, grasp it, reach a random goal position, and open the fingers to place the object. To assess the model's capabilities in dynamic environments, at each trial a velocity with random direction is assigned to the object, so that the grasping action might fail if the object moves too fast.

### 2.1 Flexible intentions for dynamic multi-step tasks

Whether an agent maintains a simple continuous generative model or makes use of a higher-level discrete model for planning, to operate in realistic scenarios it needs the notion of *environmental beliefs*. Maintaining additional beliefs over the environment has been used for actively inferring object-centric representations [41], and we here turn to the problem of tackling dynamic tasks that comprise multiple steps or composite movements. If the agent's goal is to reach a moving object, it must maintain not only a best guess about its hand position, but also about the object position [42]. Designing a dynamics function with an attractor toward a static object position is not helpful in dynamic contexts, and directly providing exteroceptive observations of the object is not biologically plausible, assuming that a hierarchical generative model infers through exchange of local messages.

Formally, we consider hidden states or beliefs $\boldsymbol{\mu}$ about different environmental entities operating in the same domain, i.e., $\boldsymbol{\mu} = [\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N]$, where $N$ is the number of entities. We assume that these beliefs generate observations $\boldsymbol{y}$ in parallel, through a likelihood function $\boldsymbol{g}(\boldsymbol{\mu})$ that maintains the same factorization:

$$\boldsymbol{y} = [\boldsymbol{y}_1 \quad \dots \quad \boldsymbol{y}_N] = [\boldsymbol{g}_1(\boldsymbol{\mu}_1) + \boldsymbol{w}_1 \quad \dots \quad \boldsymbol{g}_N(\boldsymbol{\mu}_N) + \boldsymbol{w}_N] \tag{1}$$

where the letter $w$ indicates a (Gaussian) noise term. Each belief can then be dynamically inferred by inverting this function:

$$\partial_\mu \boldsymbol{y} = [\partial_{\mu_1} \boldsymbol{g}_1 \quad \dots \quad \partial_{\mu_N} \boldsymbol{g}_N] \tag{2}$$

At this point, the dynamic nature of a task can be tackled through the definition of *intentions* [42], defined as functions that generate a possible future state (e.g., about the agent's hand) from the current belief:

$$\boldsymbol{\mu}^* = \boldsymbol{i}(\boldsymbol{\mu}) = \boldsymbol{H}\boldsymbol{\mu} + \boldsymbol{h} \tag{3}$$

Note that differently from the likelihood function $\boldsymbol{g}(\boldsymbol{\mu})$, the function $\boldsymbol{i}(\boldsymbol{\mu})$ manipulates and combines all the environmental entities. In particular, the matrix $\boldsymbol{H}$ performs a linear transformation of the beliefs, affording dynamic behavior since every entity is constantly updated through the observations. On the other hand, the vector $\boldsymbol{h}$ imposes a static configuration, e.g., a fixed position to reach. To clarify how such flexible intentions work, it is useful to model the pick-and-place operation considered in the Introduction. To focus on the control aspects, we assume a simulated agent endowed with the following simplified sensory modalities: (i) a proprioceptive observation of the arm's joint angles (with dimension $n_p = 8$); (ii) a visual observation encoding the positions of hand, object, and goal (with dimension $n_v = 3$); (iii) a tactile observation:

$$\boldsymbol{y} = \{\boldsymbol{y}_p, \boldsymbol{y}_v, y_t\} \tag{4}$$

Visual features may be inferred from a rich visual input [42], but for simplicity we directly provide the 2D Cartesian coordinates of the hand, object, and goal, i.e., $\boldsymbol{y}_v = [\boldsymbol{y}_{v,a}, \boldsymbol{y}_{v,o}, \boldsymbol{y}_{v,g}]$. Tactile observations inform the agent through a Boolean function whether all the fingers – the last four joints – touch an external object. Here, we only consider the Cartesian position of the last level (i.e., the hand), although a more realistic scenario should comprise information about all intermediate limb positions, which may be used to infer the correct posture from exteroceptive sensations only [43].

Although limb trajectories are generated by a continuous model encoding *intrinsic* kinematic information (e.g., joint angles), the movements that the agent wants to realize are usually expressed in an *extrinsic* domain (e.g., 3D visual space). These two continuous modalities are needed in pick-and-place operations because reaching movements are described in an extrinsic reference frame, while opening and closing the hand are operations better understood from an intrinsic perspective. We thus consider beliefs in both intrinsic ($\boldsymbol{\mu}_i$) and extrinsic ($\boldsymbol{\mu}_e$) domains. Note that it is convenient to express in these two domains also the objects that the agent interacts with. In this way, for each object the agent can automatically infer a possible joint configuration, which can be further restricted by some priors (e.g., an affordance depending on a particular grip). The reason for this will be clear from an example in the Supplementary Materials. We further assume that the agent maintains a belief over limb lengths $\boldsymbol{\mu}_l$ in order to compute the extrinsic Cartesian coordinates of the hand and every object. Here, this belief is not updated, although it is possible to learn it through inference [43, 44]. Finally, we also keep a belief over tactile observations $\mu_t$. Summing up:

$$\boldsymbol{\mu} = \{\boldsymbol{\mu}_i, \boldsymbol{\mu}_e, \boldsymbol{\mu}_l, \mu_t\} \tag{5}$$

where the first two beliefs comprise components representing the arm, object, and goal, reflecting the decomposition of the visual sensory signal:

$$\boldsymbol{\mu}_i = [\boldsymbol{\mu}_{i,a} \quad \boldsymbol{\mu}_{i,o} \quad \boldsymbol{\mu}_{i,g}] \qquad\qquad \boldsymbol{\mu}_e = [\boldsymbol{\mu}_{e,a} \quad \boldsymbol{\mu}_{e,o} \quad \boldsymbol{\mu}_{e,g}] \tag{6}$$

Since joint angles generate Cartesian coordinates with a one-to-one mapping, it is natural to place the intrinsic belief at the top of this hierarchy, and define the following likelihood function $\boldsymbol{g}_e$:

$$\begin{aligned} \boldsymbol{g}_e(\boldsymbol{\mu}_i) &= [\boldsymbol{T}(\boldsymbol{\mu}_{i,a}, \boldsymbol{\mu}_l) \quad \boldsymbol{T}(\boldsymbol{\mu}_{i,o}, \boldsymbol{\mu}_l) \quad \boldsymbol{T}(\boldsymbol{\mu}_{i,g}, \boldsymbol{\mu}_l)] + \boldsymbol{w}_e \\ \boldsymbol{T}(\boldsymbol{\theta}, \boldsymbol{l}) &= \begin{bmatrix} l_1 c_1 + l_2 c_{12} + l_3 c_{123} + (l_4 + l^*) c_{1234} \\ l_1 s_1 + l_2 s_{12} + l_3 s_{123} + (l_4 + l^*) s_{1234} \end{bmatrix} \end{aligned} \tag{7}$$

where $\boldsymbol{T}(\boldsymbol{\theta}, \boldsymbol{l})$ is the *forward kinematics*. In this way, the intrinsic belief can be found by inferring the most likely kinematic configuration that may have generated the extrinsic belief [43]; thus, an inverse model within the dynamics function is not necessary. Here, $l_n$ is the length of the $n$-th segment, and we used a compact notation to indicate the sine and cosine of the sum of angles. Since we are not interested in the fingers' positions during the reaching task, the kinematic likelihood function only computes the Cartesian position of the hand, which is found by extending the length of the last limb by a grasping distance $l^*$. Proprioceptive predictions are generated through a mapping $\boldsymbol{g}_p$ that extracts the arm's joint angles from the intrinsic belief, while exteroceptive sensations – here, position and touch – are generated by an identity function:

The belief dynamics can take into account many factors including friction, gravity, etc. – but in the following we assume that it depends on a combination of the intentions defined above. For the pick-and-place task considered here, it is useful to define an intention for each step: (i) reach the object; (ii) reach the goal position; (iii) close the fingers to grasp the object; (iv) open the fingers to release the object. The first two intentions can be realized in an intrinsic reference frame that sets the arm's joint angles equal to the inferred object/goal configuration

$$\begin{aligned} \boldsymbol{\mu}^*_{i,0} &= \boldsymbol{i}_{i,0}(\boldsymbol{\mu}_i) = \boldsymbol{H}_{i,0}\boldsymbol{\mu}_i + \boldsymbol{h}_{i,0} = [\boldsymbol{\mu}_{i,o} \quad \boldsymbol{\mu}_{i,o} \quad \boldsymbol{\mu}_{i,g}] \\ \boldsymbol{\mu}^*_{i,1} &= \boldsymbol{i}_{i,1}(\boldsymbol{\mu}_i) = \boldsymbol{H}_{i,1}\boldsymbol{\mu}_i + \boldsymbol{h}_{i,1} = [\boldsymbol{\mu}_{i,g} \quad \boldsymbol{\mu}_{i,o} \quad \boldsymbol{\mu}_{i,g}] \end{aligned} \tag{8}$$

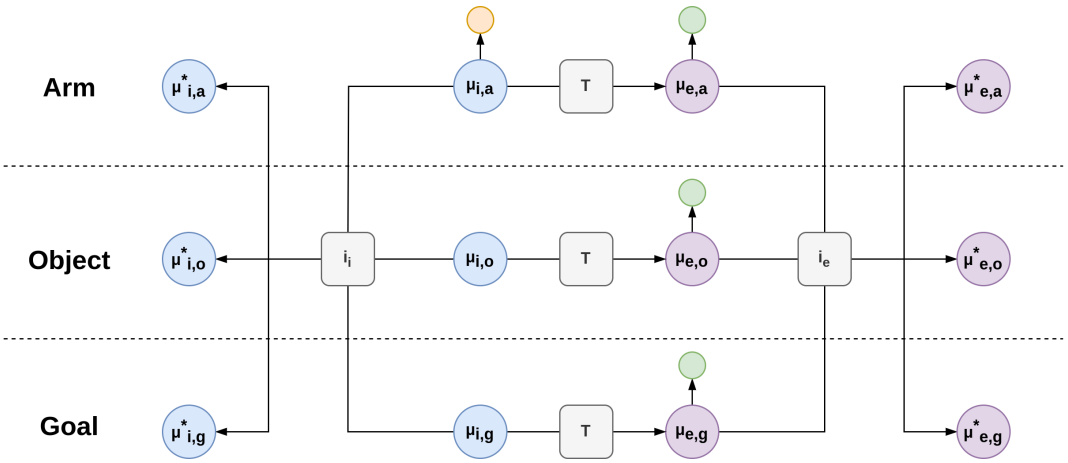



Figure 1: Graphical representation of environmental beliefs and intentions. Intrinsic and extrinsic beliefs are linked by a function $\boldsymbol{T}$ computing forward kinematics. Both beliefs have components representing the arm, an object, and a goal position. Orange and green circles represent proprioceptive and visual observations, respectively (only the arm component generates the former). Note that future states are computed through intentions $\boldsymbol{i}_i$ and $\boldsymbol{i}_e$ that manipulate and combine all the environmental entities from a given domain.

with the help of the following linear transformation matrices:

$$\boldsymbol{H}_{i,0} = \begin{bmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \qquad \boldsymbol{H}_{i,1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$
$$\boldsymbol{h}_{i,0} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad \boldsymbol{h}_{i,1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{9}$$

Intentions with a similar form can be interchangeably defined at the extrinsic level: however, in this case the prediction errors of the extrinsic belief have to be backpropagated to the intrinsic level, thus resulting in slower dynamics compared to intrinsic intentions directly operating on the arm's kinematic configuration. The grasping/placing intentions are instead accomplished by generating a future intrinsic belief with the fingers in a fixed closed/open configuration:

$$\begin{aligned} \boldsymbol{\mu}^*_{i,2} &= \boldsymbol{i}_{i,2}(\boldsymbol{\mu}_i) = \boldsymbol{H}_{i,2}\boldsymbol{\mu}_i + \boldsymbol{h}_{i,2} = \begin{bmatrix} \boldsymbol{\mu}^*_{i,2,a} & \boldsymbol{\mu}_{i,o} & \boldsymbol{\mu}_{i,h} \end{bmatrix} \\ \boldsymbol{\mu}^*_{i,3} &= \boldsymbol{i}_{i,3}(\boldsymbol{\mu}_i) = \boldsymbol{H}_{i,3}\boldsymbol{\mu}_i + \boldsymbol{h}_{i,3} = \begin{bmatrix} \boldsymbol{\mu}^*_{i,3,a} & \boldsymbol{\mu}_{i,o} & \boldsymbol{\mu}_{i,h} \end{bmatrix} \end{aligned} \tag{10}$$

Here, $\boldsymbol{H}_{i,2}$ and $\boldsymbol{H}_{i,3}$ ensure that the first 4 components of the arm (thus, excluding the fingers) maintain the current configuration:

$$\boldsymbol{H}_{i,2} = \boldsymbol{H}_{i,3} = \begin{bmatrix} \boldsymbol{H}_{i,a} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \qquad \boldsymbol{H}_{i,a} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{11}$$

while $\boldsymbol{h}_{i,2,a}$ and $\boldsymbol{h}_{i,3,a}$ impose to the fingers the correct closed/open angles $\theta_c$ and $\theta_o$:

$$\begin{aligned} \boldsymbol{h}_{i,2} &= \begin{bmatrix} \boldsymbol{h}_{i,2,a} & \mathbf{0} & \mathbf{0} \end{bmatrix} & \boldsymbol{h}_{i,2,a} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -\theta_c & \theta_c \end{bmatrix} \\ \boldsymbol{h}_{i,3} &= \begin{bmatrix} \boldsymbol{h}_{i,3,a} & \mathbf{0} & \mathbf{0} \end{bmatrix} & \boldsymbol{h}_{i,3,a} &= \begin{bmatrix} 0 & 0 & 0 & 0 & \theta_o & -\theta_o & 0 & 0 \end{bmatrix} \end{aligned} \tag{12}$$

In summary, we exemplified a context with two sets of competing intentions. To make the picture clearer, Figure 1 provides a graphical representation of the model linking environmental beliefs and flexible intentions. In the following, we turn to the definition of the two control methods, explaining how these intentions are used in practice. What differs between a hybrid and a continuous-only model is where they are embedded and how they affect the system dynamics.

### 2.2 Hybrid models and discrete goals

Designing a discrete model involves the specification of discrete states $\boldsymbol{s}$, a likelihood matrix $\boldsymbol{A}$, a transition matrix $\boldsymbol{B}$, and actions $\boldsymbol{u}$. In the task considered, the discrete hidden states encode: (i) whether the hand and object are in the same position, at different positions, or both at the goal position; and (ii) whether the hand is open, closed or has grasped the object. In total, these factors combine in 9 possible process states. There are two likelihood matrices: (i) a matrix $\boldsymbol{A}$ performs a simple identity mapping and is used to obtain the discrete outcomes for the continuous model; and (ii) a matrix $\boldsymbol{A}_t$ returns a discrete tactile observation $\boldsymbol{o}_t$ – encoded by a Bernoulli distribution – signaling whether or not the object is grasped (here, the continuous observation $y_t$ is not used). The transition matrix $\boldsymbol{B}$ is defined such that the object can be grasped only when both hand and object are in the same position. Note that every state has a certain probability – depending on the continuous sampling time – that the transition might fail. Finally, the discrete actions $\boldsymbol{u}$ correspond to the intentions defined above, with the addition of a *stay* action. The trial is considered successful only if the agent has placed the object at the goal position and the hand is open.

For the discrete and continuous models to communicate, each discrete outcome has to be associated with a reduced continuous prior. The latter is usually static, hence in order to deal with a dynamic environment we let it depend on a combination of the intentions defined before. This has the advantage that if the reduced priors are generated at each discrete time step, a correct configuration over the hidden states can be imposed even if an object is moving – whose position will be dynamically inferred by the continuous model. However, note that a more realistic design would be to link the reduced priors to the hidden causes and let the latter encode the intentions [45]. Since the agent maintains both intrinsic and extrinsic beliefs, a single discrete outcome $o_{\tau,m}$ generates two different sets of reduced priors, $\boldsymbol{\eta}_{i,m}$ and $\boldsymbol{\eta}_{e,m}$. For example, if the hand is open and at the object position, the reduced priors are:

$$\boldsymbol{\eta}_{i,0} = \boldsymbol{i}_{i,2}(\boldsymbol{i}_{i,0}(\boldsymbol{\mu}_i)) \qquad \boldsymbol{\eta}_{e,0} = \boldsymbol{i}_{e,0}(\boldsymbol{\mu}_e) \tag{13}$$

Note that some outcomes generate the same reduced priors (e.g., the open hand condition does not affect the extrinsic prior), or do not impose any bias over the hidden states (e.g., if the hand and the object are at different positions, the mapping of the extrinsic level reduces to an identity and the corresponding reduced prior will be equal to the posterior). Through a Bayesian Model Average (BMA), the full priors $\boldsymbol{\eta}_i$ and $\boldsymbol{\eta}_e$ are then computed, which act over the continuous models through the prior errors $\boldsymbol{\varepsilon}_{\eta,i}$ and $\boldsymbol{\varepsilon}_{\eta,e}$:

$$\begin{aligned} \boldsymbol{\eta}_i &= \sum_m o_{\tau,m} \cdot \boldsymbol{\eta}_{i,m} & \boldsymbol{\varepsilon}_{\eta,i} &= \boldsymbol{\mu}_i - \boldsymbol{\eta}_i \\ \boldsymbol{\eta}_e &= \sum_m o_{\tau,m} \cdot \boldsymbol{\eta}_{e,m} & \boldsymbol{\varepsilon}_{\eta,e} &= \boldsymbol{\mu}_e - \boldsymbol{\eta}_e \end{aligned} \tag{14}$$

The updates for the hidden states are finally:

$$\begin{aligned} \dot{\boldsymbol{\mu}}_i &= -\partial_{\boldsymbol{\mu}_i}\mathcal{F} = \partial \boldsymbol{g}_p^T \pi_p \boldsymbol{\varepsilon}_p + \partial \boldsymbol{g}_e^T \pi_e \boldsymbol{\varepsilon}_e - \pi_{\eta,i}\boldsymbol{\varepsilon}_{\eta,i} \\ \dot{\boldsymbol{\mu}}_e &= -\partial_{\boldsymbol{\mu}_e}\mathcal{F} = \partial \boldsymbol{g}_v^T \pi_v \boldsymbol{\varepsilon}_v - \pi_e \boldsymbol{\varepsilon}_e - \pi_{\eta,e}\boldsymbol{\varepsilon}_{\eta,e} \end{aligned} \tag{15}$$

where $\mathcal{F}$ is the variational free energy of the continuous models, while $\boldsymbol{\varepsilon}_p$, $\boldsymbol{\varepsilon}_e$, $\boldsymbol{\varepsilon}_v$ are respectively the proprioceptive, extrinsic, and visual prediction errors, and we indicated with the letter $\pi$ the precisions of their likelihood. Note that both updates have a similar form, with the only differences lying in the type of sensory observations, and in the direction of the extrinsic prediction error.

On the other side, ascending messages are computed by comparing, for a period $T$, the models of both intrinsic and extrinsic posteriors through the reduced free energies $\tilde{\mathcal{F}}_m$ (see the Methods section for a detailed explanation):

$$\tilde{\mathcal{F}}_m = -\ln o_{\tau,m} - \delta_i \sum_0^T \mathcal{L}_{m,i} dt - \delta_e \sum_0^T \mathcal{L}_{m,e} dt \tag{16}$$

where $\mathcal{L}_{m,i}$ and $\mathcal{L}_{m,e}$ are the log-evidences of the intrinsic and extrinsic beliefs, respectively weighted by the gains $\delta_i$ and $\delta_e$. The latter can be seen as quantities that modulate the magnitude of all precisions involving a specific domain. Here, they are kept fixed throughout the tasks but it is assumed that higher values result in faster reaction times and evidence accumulation during perceptual decision-making [46]. Note that a long enough *time window* $T$ is needed to accumulate evidence from the lower modalities and plan the next primitive movement, as will be clear later. At this point, the current state is inferred by combining the likelihoods of both the ascending message and the tactile observation:

$$\boldsymbol{s}_{\pi,\tau} = \sigma(\ln \boldsymbol{B}_{\pi,\tau-1}\boldsymbol{s}_{\tau-1} + \ln \boldsymbol{A}^T \boldsymbol{r}_\tau + \ln \boldsymbol{A}_t^T \boldsymbol{o}_t) \tag{17}$$

In summary, the discrete model first computes policy probabilities depending on prior preferences, and then averages them to obtain a discrete outcome. The latter is used to compute – through flexible intentions – the reduced priors,

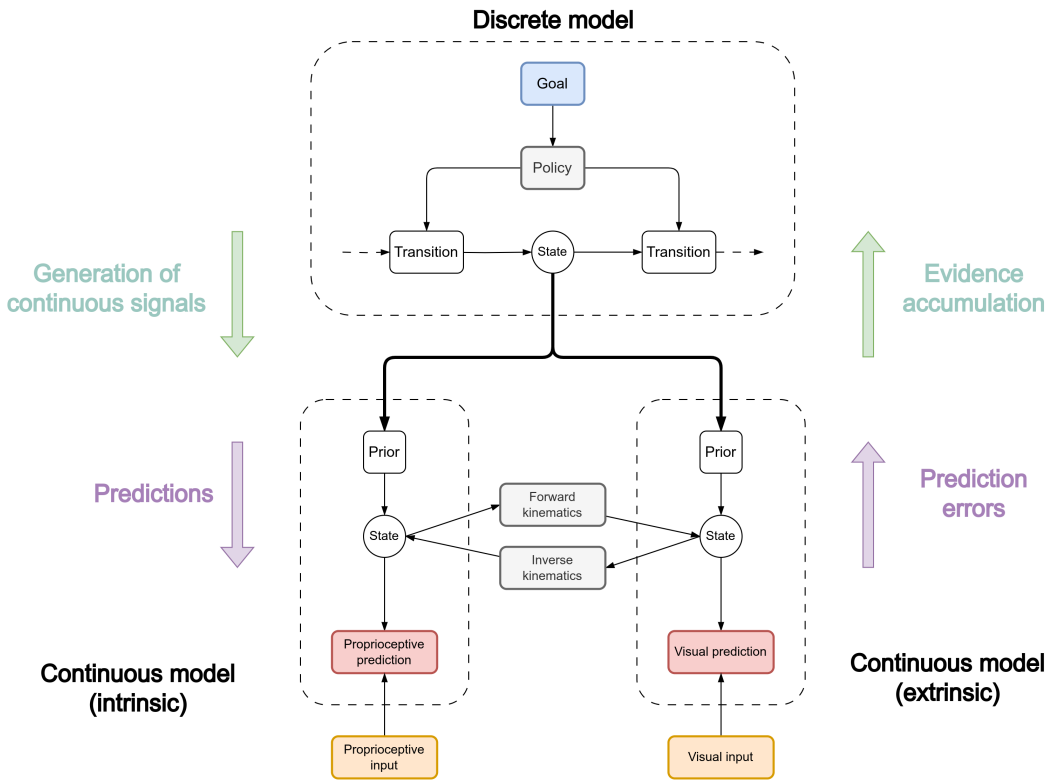



Figure 2: Graphical representation of the first cognitive phase. In the first trials of every novel task, the state transitions generally encode elementary actions, such as reaching, grasping, and releasing. The cognitive effort needed by high-level processes to learn the correct sequence of elementary actions translates to the task dynamics being only visible in the discrete model, while the continuous levels only receive a prior representation (e.g., target states) and their job is mostly to realize the corresponding trajectories. This is in line with the hypothesis that in the first phase of motor learning a high cortical activity is recorded, especially in anterior (e.g., prefrontal) regions.

which are further weighted to get a final configuration that the continuous models have to realize. This configuration acts as a prior over the 0th-order (i.e., position) hidden states, and the belief update breaks down to the computation of sensory gradients and prediction errors. Note that the extrinsic belief $\boldsymbol{\mu}_e$ can be biased through two different pathways: directly from the reduced prior of the discrete model, or indirectly through forward kinematics of the intrinsic belief $\boldsymbol{\mu}_i$. Correspondingly, the latter can be changed either by imposing a particular configuration in the intrinsic domain through the discrete model, or by the backward pathway of the extrinsic belief, through computation of the gradient $\partial \boldsymbol{g}_e$ [43].

What are the implications from a biological perspective? It has been hypothesized that motor skill learning is roughly composed of three different phases [47]: a first cognitive phase where movements are inefficient and one has to consciously explore different strategies to achieve a goal; an associative stage where the learner becomes more proficient and transitions between movements are more fluid; a final autonomous phase where the learner has reached almost complete autonomy from conscious task-specific thinking and can move without cognitive effort. Although the reality is likely to be far more complex than this simple decomposition, with different degrees of proficiency corresponding to specific neural processing, it may be useful for grasping and analyzing the key principles that underlie the change in neural activity about task specialization. In this view, active inference provides interesting insights about this process because: (i) it builds upon an internal generative model of the task dynamics, which can be constantly refined by experience; (ii) it assumes that the whole information processing starts and unfolds through the same principle of prediction error minimization, either with continuous or discrete representations. Under this perspective, the hybrid model described can be compared with the first cognitive phase, as shown in Figure 2.

### 2.3 Continuous models and flexible intentions

Whether or not a discrete model is necessary depends on the nature of the task, even when considering movements composed of different steps. In some cases, an a-priori-defined sequence of actions (i.e., a habitual behavior) is all that

is needed to solve a particular task, e.g., in rhythmic movements or, as in this case, a simple pick-and-place operation. Such scenarios may present little uncertainty about the order of elementary actions that does not necessarily involve repeated online decision-making.

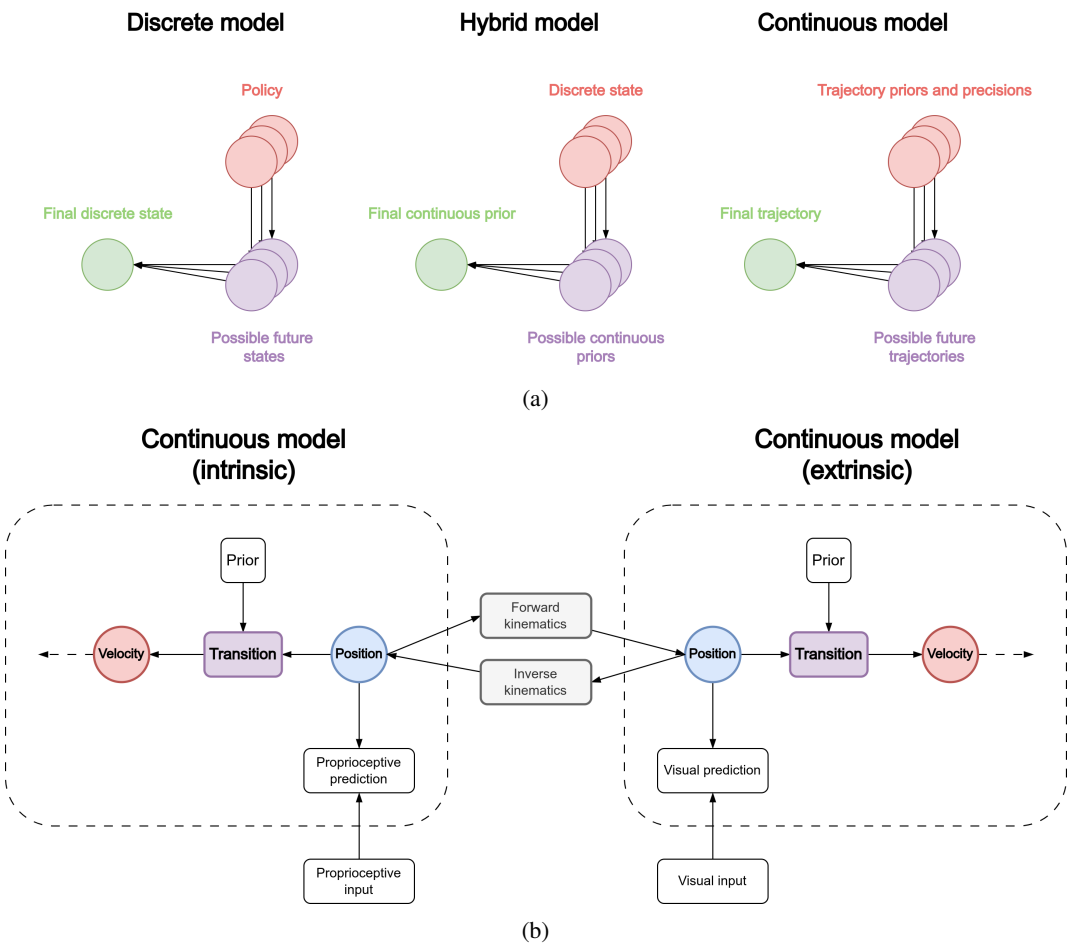



Figure 3: (A) Generation of composite states in discrete, hybrid, and continuous models. (B) Graphical representation of the third autonomous phase. The process of task specialization is complete. The discrete model – not shown here – has learned a representation abstract enough to include all the four main movements of the pick-and-place operation, and its role is only to command its execution and accumulate evidence from both modalities to infer when it has been completed.

But how can transitions between continuous trajectories be encoded in practice? To elucidate, we reveal an interesting parallelism between high-level and low-level processes, shown in Figure 3a. In the discrete model, states are conditioned on every policy that the agent can take in a specific instant through corresponding probability distributions, and the overall state is found by averaging the expected states from all transitions. The same technique is applied to hybrid models, where the probability of every possible discrete outcome is combined with a continuous state to compute the average signal that will bias the lower levels. Along the same line, we propose that composite continuous dynamics can be generated by averaging the results from independent contributions, each encoding a primitive trajectory. Specifically, we resort to the notion of *hidden causes* (see the Methods section) to encode the agent's intentions, i.e., $\boldsymbol{\nu}_i = [\nu_{i,1}, \ldots, \nu_{i,J}]$ and $\boldsymbol{\nu}_e = [\nu_{e,1}, \ldots, \nu_{e,K}]$, where $J$ and $K$ are the number of intrinsic and extrinsic intentions. Next, we factorize the prior distribution over the generalized beliefs $\tilde{\boldsymbol{\mu}}_i$ and $\tilde{\boldsymbol{\mu}}_e$ (in this case, up to the 1st order) into different probability distributions for each hidden cause:

$$p(\tilde{\boldsymbol{\mu}}_i) = p(\boldsymbol{\mu}_i) \prod_j^J p(\boldsymbol{\mu}'_{i,j} | \boldsymbol{\mu}_i) \qquad p(\tilde{\boldsymbol{\mu}}_e) = p(\boldsymbol{\mu}_e) \prod_k^K p(\boldsymbol{\mu}'_{e,k} | \boldsymbol{\mu}_e) \tag{18}$$

where:

$$p(\boldsymbol{\mu}'_{i,j}|\boldsymbol{\mu}_i) = \mathcal{N}(\boldsymbol{\mu}'_{i,j}|\boldsymbol{f}_{i,j}(\boldsymbol{\mu}_i, \nu_{i,j}), \pi^{-1}_{\mu,i,j}) \qquad p(\boldsymbol{\mu}'_{e,k}|\boldsymbol{\mu}_e) = \mathcal{N}(\boldsymbol{\mu}'_{e,k}|\boldsymbol{f}_{e,k}(\boldsymbol{\mu}_e, \nu_{e,j}), \pi^{-1}_{\mu,e,k}) \tag{19}$$

and $\pi_{\mu,i,j}$ and $\pi_{\mu,e,k}$ are the corresponding precisions. Each function acts as an attractive force proportional to the error between the current and desired beliefs:

$$\begin{aligned} \boldsymbol{f}_{i,j}(\boldsymbol{\mu}_i, \nu_{i,j}) &= \nu_{i,j}\boldsymbol{e}_{i,j} & \boldsymbol{e}_{i,j} &= \boldsymbol{\mu}^*_{i,j} - \boldsymbol{\mu}_i \\ \boldsymbol{f}_{e,k}(\boldsymbol{\mu}_e, \nu_{e,j}) &= \nu_{e,j}\boldsymbol{e}_{e,k} & \boldsymbol{e}_{e,k} &= \boldsymbol{\mu}^*_{e,k} - \boldsymbol{\mu}_e \end{aligned} \tag{20}$$

where $\boldsymbol{\mu}^*_{i,j}$ and $\boldsymbol{\mu}^*_{e,k}$ are the future states computed through the intentions defined above. Note that the hidden causes can either act as attractor gains, or specify the relative strength of one state over the other [42, 48]; as a result, modulation of the hidden causes achieves a fixed multi-step behavior. In fact, since the belief dynamics already store and embed every future goal, the belief will follow the contribution of all active states. In particular, we want the following behavior: (i) until the hand has not reached the object, execute only $\boldsymbol{i}_{i,0}$ (i.e., reach the object) and $\boldsymbol{i}_{i,3}$ (i.e., open the hand); (ii) as soon as the hand reaches the object, $\boldsymbol{i}_{i,2}$ should start (i.e., close the hand); (iii) when the agent has grasped the object, the object-reaching intention should be replaced by the $\boldsymbol{i}_{i,1}$ (i.e., move the hand toward the goal position along with the object); (iv) when the agent has reached the goal position, execute $\boldsymbol{i}_{i,3}$ again to release the object. Note that the object-reaching intention should be active until the agent has correctly grasped the object, since the latter may be moving and the grasping intention fail. The desired behavior can be easily implemented by computing and combining Boolean functions (or sigmoid functions to have smooth transitions) of the beliefs. Having defined the dynamics prediction errors for each intention:

$$\boldsymbol{\varepsilon}_{\mu,i,j} = \boldsymbol{\mu}'_i - \boldsymbol{\mu}'_{i,j} \qquad \boldsymbol{\varepsilon}_{\mu,e,k} = \boldsymbol{\mu}'_e - \boldsymbol{\mu}'_{e,k} \tag{21}$$

the update for the intrinsic and extrinsic beliefs will be a combination of sensory prediction errors and attractor dynamics:

$$\begin{aligned} \dot{\tilde{\boldsymbol{\mu}}}_i &= \mathcal{D}\tilde{\boldsymbol{\mu}}_i - \partial_{\tilde{\boldsymbol{\mu}}_i}\mathcal{F} = \begin{bmatrix} \boldsymbol{\mu}'_i + \partial\boldsymbol{g}_p^T\pi_p\boldsymbol{\varepsilon}_p + \partial\boldsymbol{g}_e^T\pi_e\boldsymbol{\varepsilon}_e + \sum\partial\boldsymbol{f}_{i,j}^T\pi_{\mu,i,j}\boldsymbol{\varepsilon}_{\mu,i,j} \\ -\sum\pi_{\mu,i,j}\boldsymbol{\varepsilon}_{\mu,i,j} \end{bmatrix} \\ \dot{\tilde{\boldsymbol{\mu}}}_e &= \mathcal{D}\tilde{\boldsymbol{\mu}}_e - \partial_{\tilde{\boldsymbol{\mu}}_e}\mathcal{F} = \begin{bmatrix} \boldsymbol{\mu}'_e + \partial\boldsymbol{g}_v^T\pi_v\boldsymbol{\varepsilon}_v - \pi_e\boldsymbol{\varepsilon}_e + \sum\partial\boldsymbol{f}_{e,k}^T\pi_{\mu,e,k}\boldsymbol{\varepsilon}_{\mu,e,k} \\ -\sum\pi_{\mu,e,k}\boldsymbol{\varepsilon}_{\mu,e,k} \end{bmatrix} \end{aligned} \tag{22}$$

Finally, we update the beliefs over tactile sensations:

$$\dot{\boldsymbol{\mu}}_t = \partial\boldsymbol{g}_t^T\pi_t\boldsymbol{\varepsilon}_t \tag{23}$$

Figure 3b shows a graphical representation of this model, compared to the third phase of motor learning. The continuous models now encode the dynamics of the entire task in the most efficient way because the mechanism is fully autonomous and does not require the repeated activation of higher levels for planning the next action. Differently than the previous model (Figure 2), which had to encode every possible elementary movement – either intrinsic or extrinsic – the increased efficiency is also the consequence of each modality encoding its dynamics independently from the others, translating to grasping (or reaching) movements being mapped mostly into the intrinsic (or extrinsic) modality. Note in fact that the dimensions of hidden causes $J$ and $K$ generally differ. A sequence of time frames is shown in Figure 4 for a sample grasping trial with this continuous-only control. Note that both the intrinsic and extrinsic beliefs comprise estimates of the arm, the object to be grasped, and the goal position.

### 2.4 Discrete and continuous processes compared

In summary, goal-directed behavior can be achieved through specular mechanisms operating at different hierarchical locations, as shown in Figure 3a. Comparing the continuous belief update with the discrete state inference, we note a relationship between the policy-dependent $\boldsymbol{s}_{\pi,\tau}$ states computed through the transition matrices, and the trajectories $\boldsymbol{\mu}'_{i,j}$ and $\boldsymbol{\mu}'_{e,k}$ computed through the continuous intentions. While a policy-independent discrete state $\boldsymbol{s}_\tau$ is found by calculating a BMA with each policy probability $\pi_\pi$, in the continuous case the hidden states $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_e$ are computed by averaging each trajectory with the corresponding hidden causes $\nu_{i,j}$ and $\nu_{e,k}$ at each step of the task.

If we consider the hybrid model, we also note a similarity between the probability distributions of the reduced priors of Equation 50 and the probabilities of Equation 19 that compose the dynamics function. In the former case, each reduced prior is averaged through each outcome model probability $o_{\tau,m}$. In this case, however, the full prior biases the hidden states through the overall prediction error $\varepsilon_{\eta,i}$, which already contains the final configuration that the continuous model must realize until the next discrete step. In the continuous-only model, the intentions are instead used to compute

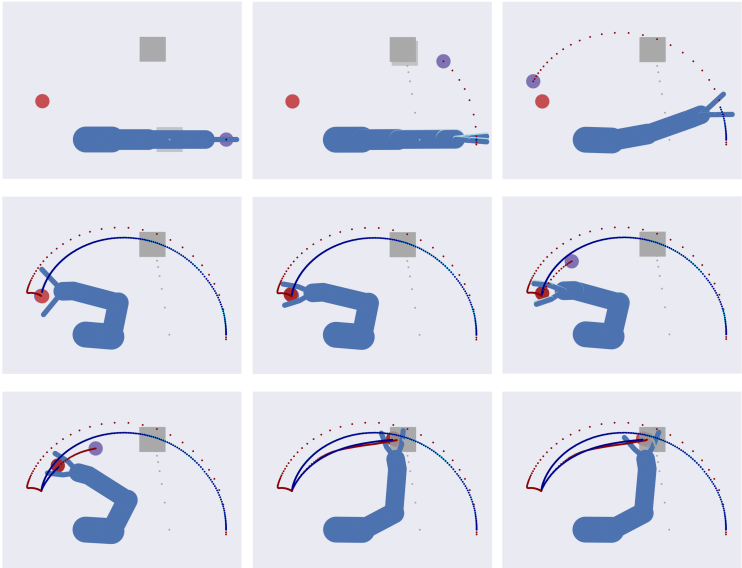


Figure 4: Sequence of time frames of the grasping task. Real and estimated object positions are represented by red and purple circles, real and estimated goal positions by grey and dark grey squares, and real and estimated limb positions in blue and cyan. Hand, object, and goal belief trajectories are represented respectively with a blue, red, and grey dotted line.

different directions of update that independently pull the belief over the hidden states. Also, the prior precision $\pi_{\eta,i}$ of Equation 15 and the dynamics precisions $\pi_{\mu,i,j}$ and $\pi_{\mu,e,k}$ of Equation 22 play the same role of encoding the strength of the intention dynamics – that steers the belief toward a desired state – with respect to the sensory likelihood – that keeps the agent close to what it is currently perceiving. Crucially, while the weighted prediction error of the hybrid model is:

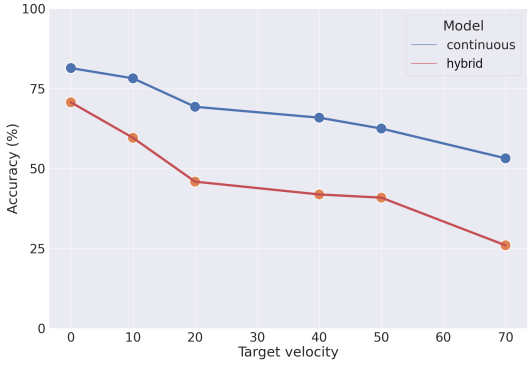
$$\pi_{\eta,i}(\boldsymbol{\mu}_i - \sum o_{\tau,m}\boldsymbol{\eta}_{i,m}) \tag{24}$$

if the dynamics precisions have the same value $\pi_{\mu,i}$, the combination of the weighted dynamics prediction error of the continuous-only model takes the form:

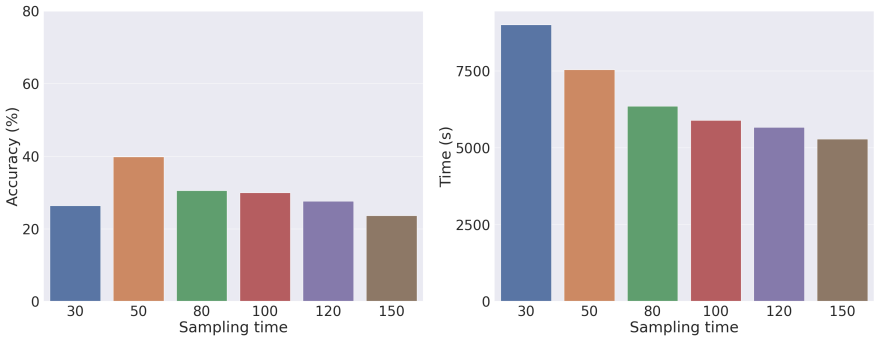
$$\boldsymbol{\mu}'_i + \pi_{\mu,i}(\boldsymbol{\mu}_i - \sum \nu_{i,j}\boldsymbol{\mu}^*_{i,j}) \tag{25}$$

However, note that encoding different precisions for each intention permits an additional modulation. In fact, in the continuous-only model the future trajectories are weighted by two different processes [45]. A *fast* process that imposes (and infers) future states based on the hidden causes $\boldsymbol{\nu}_i$ and $\boldsymbol{\nu}_e$; but also a *slow* process that learns the precisions $\pi_{\mu,i,j}$ and $\pi_{\mu,e,k}$ of every trajectory. This result lends itself to an intuitive interpretation: in active inference, a low precision in a specific sensory modality implies that it cannot be trusted and the agent should update its internal state relying on other signals. Correspondingly, a trajectory with high precision will be a good option for minimizing the prediction errors in the current context, i.e., the agent will be confident in using it for the task considered.

Figure 5a displays the capacity of the continuous and hybrid models to solve the task in dynamic conditions, using a fixed sampling time window – the period $T$ in Equation 16 – for the hybrid model. While the two models achieve somewhat similar performances in static grasping (i.e., zero object velocity), the difference increases when it comes to grasping a moving object. The performance drop of the hybrid model is caused by the amount of sampling of continuous evidence performed by the discrete part, as illustrated in Figure 5b, which reveals a tradeoff between reaction times and model complexity. On one hand, greater sampling time allows the agent to accumulate more information and hence estimate more accurately the discrete state. However, longer sampling time also implies chunking the action into a smaller number of longer primitive actions, which ultimately makes the grasping action fail since the moving object continuously changes its location during the sampling period. On the contrary, reducing the sampling time has the advantage of speeding up the agent's planning process (i.e., lower reaction times) so that it can grasp the object with greater accuracy, but only up to a point. In this case, maintaining the same performances would require longer policies due to a greater number of continuous trajectories needed for a reaching movement. At the same time, since the discrete model is activated at a higher rate for action replanning, the simulation time steadily increases too, which can be associated with higher energy demands.




(a)




(b)

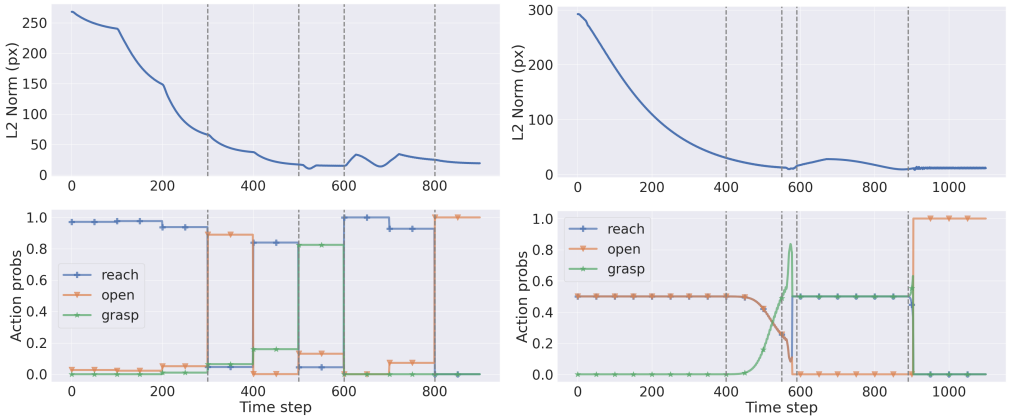
Figure 5: (A) Performance comparison between the hybrid and continuous-only models for different object velocities. Accuracy measures the number of successful trials. As the velocity of the object increases, the grasping action becomes more difficult for the hybrid model since faster response planning is needed. Instead, the continuous-only agent constantly adjusts its hidden causes based on the continuous evidence inferred at each time step. Despite using an a-priori (assumed to be learned) sequence of actions that does not permit replanning, this semi-flexible dynamics handles dynamic tasks more smoothly and efficiently. (B) Behavioral effects of the sampling time in the hybrid model with an object moving at a moderate velocity. Accuracy (left) and overall simulation time (right) are shown. Every condition is averaged over 1000 trials.

Figures 6a and 6b show the interplay between reaching and gasping actions with the hybrid model and a constant sampling time. Here, five different phases can be distinguished: a pure reaching movement, an intermediate phase when the agent slowly approaches the object and prepares the grasping action, a grasping phase, another reaching movement and, finally, the object release. The computation of the action probabilities by the discrete model thus realizes a smooth transition between reaching and grasping actions, although at this stage still encoded as separate representations.

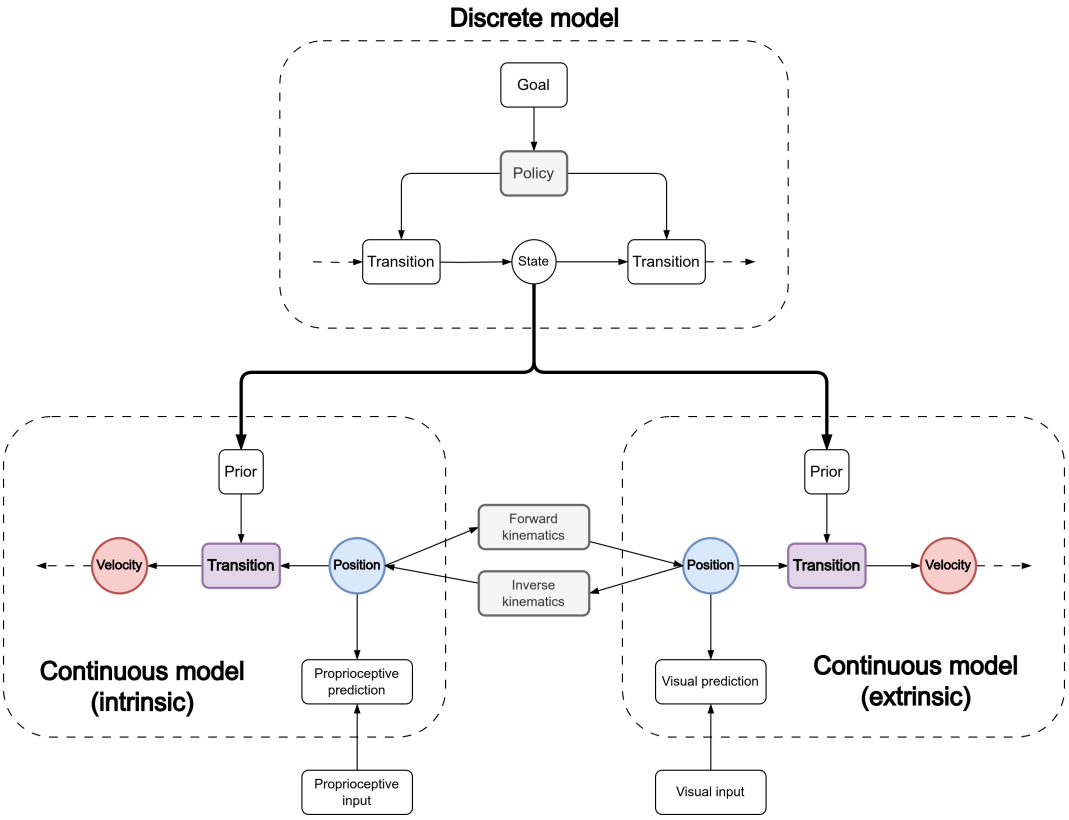
This finding might provide a clue on how discrete actions lead to the emergence of smoother composite trajectories in the continuous domain, typical of intermediate phases of motor skill learning. In the second (associative) phase, the cortical activity begins to shift toward posterior areas. The process can be modeled as in Figure 6c, with a balanced work between discrete and continuous models in which the latter start to construct their own task dynamics that will eventually lead to the model of Figure 3b. Note the differences with Figure 2, wherein the continuous models only receive a static prior resulting from a combination of all discrete states. In summary, while in the cognitive phase the task would be comprised of a reaching movement, a cognitive effort to find the next optimal action, and then a pure grasping action, in the associative phase the continuous models would handle the transition generating a fluid movement that makes the hand close as soon as it approaches the object. The discrete model could hence represent the task with a reaching-grasping action and a reaching-releasing action, with reduced computational demand.

## 3 Discussion

Understanding the mechanisms that support task specialization and motor skill learning is compelling for making advances in current robots and intelligent systems. However, a common issue is that high-level and low-level processes are often analyzed and developed with different perspectives and techniques, e.g., optimal control algorithms [49] or




(a) Action probabilities of the hybrid model. (b) Trajectory priors of the continuous model.




(c) Graphical representation of the second associative phase. Continuous dynamics are represented in generalized coordinates of motion including position, velocity, acceleration, etc., allowing one to express with increasing fidelity the true environmental dynamics. Further, the outcomes of the discrete model not only bias the 0th temporal order (i.e., position) state, but steer the whole continuous trajectory toward a particular direction. In our reach-to-grasp task, this might correspond to considering the elementary actions of reaching and grasping as single discrete states that generate composite trajectories in the continuous domain.

deep neural networks [50]. Instead, the flexibility and robustness behind human and animal motor systems reside in the fact that the learning of a new skill is a unified process that initially involves prevalent activation of prefrontal brain areas, and gradually shifts toward posterior and subcortical regions [12]. Indeed, constantly relying on the whole hierarchy is computationally demanding, and high-level areas are unable to track environmental changes rapidly. Highly dynamic tasks can be therefore solved more efficiently by offloading the learned action transitions and policies to lower, faster hierarchical layers operating in continuous domains [51, 12]. The Posterior Parietal Cortex is known to encode multiple goals in parallel during sequences of actions as in multi-step reaching, even when there is a considerable delay

between goal states [52], and evidence suggests that rhythmic and repetitive movements do not involve the continuous activation of prefrontal areas but only rely on sensorimotor circuits [53].

We propose that the computations of hybrid and continuous models of active inference can be compared to different phases of motor skill learning. When the agent is required to interact with a novel situation, high-level planning is essential because the low levels alone are unable to minimize the generated prediction errors, which then climb up the cortical hierarchy. As the agent practices the task and learns composite state transitions that already take into account environmental uncertainty and possible dynamic elements, the prediction errors arising during the unfolding of the task can be explained away more efficiently by lower-level specialized neurons, and repeated calling of high levels ceases. Even when using elementary discrete actions, a composite movement corresponding to an approaching phase between reaching and grasping naturally arises from the BMR because the continuous evidence accumulation provides a smooth transition between the probabilities over the discrete hidden states, as shown in Figures 6a and 6b. We therefore propose that the dynamics of a continuous model may have a specular structure to its discrete counterpart – i.e., the final trajectory is generated by weighting independent distributions. As a result, the composite movement might be embedded into the continuous dynamics with two parallel processes: while a discrete model can rapidly impose and infer specific trajectories, their precisions are additionally adapted through the same mechanisms of free energy minimization, so that the system will score how well an action can perform. In other words, we propose that these precisions act not just as modulatory signals but represent the agent's confidence [42] about the current state of the task, in a specular manner to sensory precisions. The higher the precision of a specific trajectory, the more useful it is in explaining the agent's final goal; on the other hand, a low trajectory precision means that the agent is not confident about its selection for a particular context.

Further, the results in Figure 5b show an interesting tradeoff between reaction times, planning capabilities, and computational demand. If the environment changes with high frequency, occasional calling of the discrete model does not permit to respond in time to novel sensory observations. On the other hand, repeated action replanning is counterproductive beyond a certain limit since, although allowing the agent to react more rapidly to environmental changes, it increases the effort of the discrete model to find the correct policy, as a higher number of discrete steps are needed. Instead, a continuous-only model achieves the best performances in less time and with small resource expenditure, but at the inability to perform replanning if the task goes into unexpected tracks. As shown in the Methods section, the presented work also contains two novelties with respect to state-of-the-art hybrid models: first, the reduced priors are generated dynamically at the beginning of each discrete step through the specification of flexible intentions – which roughly correspond to the discrete actions – letting the agent able to operate in dynamic environments. Second, the discrete model can impose priors and accumulate evidence over multiple continuous modalities, e.g., intrinsic and extrinsic. This has the advantage that one can realize more complex goals (e.g., reaching with specific affordances, as shown in the Supplementary Materials) and has access to more information for inferring the combined discrete state (e.g., extrinsic for the position, and intrinsic for the hand status).

Last, the analogies that we presented between discrete, mixed, and continuous models may shed light on the information processing within cortical regions. It has been hypothesized that the cortex works within a system of discrete models of increasing hierarchical depth (e.g., motor, associative, limbic) [11], while the interface with the continuous models that receive sensory observations is supposed to be achieved through subcortical structures, such as the superior colliculus [35] or the thalamus [34]. Since deeper cortical layers are found to encode increasingly discretized representations [18], the latter could also be the consequence of the invariability of neural activity that results during inference when the hierarchy is spatially deep. In this view, whether it makes more sense to consider a specific layer as encoding a discrete or continuous representation – cortical computations have been also simulated by continuous models [54, 55], but as well noted in [18], even the first computations of neural processing could be viewed as dealing with lots of little categories – from a high-level perspective a gradual transition between the two modalities could occur. Therefore, since all regions eventually have to deal with continuous signals, looking for analogies between the two processes might be helpful (e.g., the policy optimization of discrete models is generally compared to corticostriatal pathways [34], but reward-related activity has been recorded also in the Primary Visual Cortex V1 [56]).

Overall, these results may pave the way for explaining how task specialization could occur in biologically plausible ways. However, in this preliminary study, the different phases were analyzed separately to show the relationships between policies and dynamic trajectories. Further studies are needed to understand how the generation of prediction errors throughout the hierarchy leads to the adaptation from discrete to continuous processes. Two promising directions would be to implement deep hierarchical models [43], and to let priors and precisions of the continuous dynamics – which in the simulations were kept fixed throughout the task – adapt to the prediction errors, thus allowing self-modeling based on the current task and environmental state. Finally, future works will assess the capabilities of the hybrid active inference model to simulate human movements; for instance, encoding the object dynamics along with the agent's desires would possibly account for anticipatory effects, thus simulating more closely kinematic experiments where catching moving objects is required.

# 4 Methods

## 4.1 Perception, control, and the variational free energy

Active inference assumes that organisms perceive the environment relying on an internal generative model that is constructed by inferring how external causes produce sensations in the real generative process [17, 20]. In the continuous domain, this process is usually factorized into probability distributions over hidden causes $\tilde{\boldsymbol{v}}$, hidden states $\tilde{\boldsymbol{x}}$ and observable outcomes $\tilde{\boldsymbol{y}}$:

$$p(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}, \tilde{\boldsymbol{y}}) = p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}})p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}})p(\tilde{\boldsymbol{v}}) \tag{26}$$

These distributions are approximated by Gaussian functions:

$$\begin{aligned} p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}), \tilde{\boldsymbol{\Pi}}_y^{-1}) \\ p(\mathcal{D}\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}), \tilde{\boldsymbol{\Pi}}_x^{-1}) \\ p(\tilde{\boldsymbol{v}}|\boldsymbol{\eta}) &= \mathcal{N}(\boldsymbol{\eta}, \tilde{\boldsymbol{\Pi}}_v^{-1}) \end{aligned} \tag{27}$$

corresponding to the following non-linear stochastic equations that represent how the environment evolves:

$$\begin{aligned} \tilde{\boldsymbol{y}} &= \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{x}}) + \boldsymbol{w}_y \\ \mathcal{D}\tilde{\boldsymbol{x}} &= \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{v}}) + \boldsymbol{w}_x \end{aligned} \tag{28}$$

Here, $\mathcal{D}$ is the differential shift operator, the symbol $\sim$ indicates variables encoded in generalized coordinates representing their instantaneous trajectories, and we indicated with the symbol $\tilde{\boldsymbol{\Pi}}$ the generalized precisions of the distributions.

Directly evaluating the posterior $p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})$ is intractable since it involves the computation of the inaccessible marginal $p(\tilde{\boldsymbol{y}})$. The variational solution is to approximate the posteriors with tractable distributions, e.g., Gaussians:

$$\begin{aligned} q(\tilde{\boldsymbol{x}}) &= \mathcal{N}(\tilde{\boldsymbol{x}}; \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{P}}_x^{-1}) \\ q(\tilde{\boldsymbol{v}}) &= \mathcal{N}(\tilde{\boldsymbol{v}}; \tilde{\boldsymbol{\nu}}, \tilde{\boldsymbol{P}}_v^{-1}) \end{aligned} \tag{29}$$

where the parameters $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\nu}}$ are called beliefs over hidden states and hidden causes, while $\tilde{\boldsymbol{P}}_x$ and $\tilde{\boldsymbol{P}}_v$ are their precisions. Perceptual inference then turns into a minimization of the difference between the approximate and real posteriors, which can be formalized in terms of a KL divergence – equivalent to the expectation, over the approximate posterior, of the difference between the two log-probabilities:

$$D_{KL}[q(\tilde{\boldsymbol{x}})||p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})] = \sum\nolimits_x q(\tilde{\boldsymbol{x}}) \ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})} = \mathop{\mathbb{E}}_{q(\tilde{\boldsymbol{x}})}[\ln q(\tilde{\boldsymbol{x}}) - \ln p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})] \tag{30}$$

Given that the denominator $p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})$ still depends on the marginal $p(\tilde{\boldsymbol{y}})$, we express the KL divergence in terms of the log-evidence and the *Variational Free Energy* (VFE), and minimize the latter quantity instead. Given that the KL divergence is non-negative (due to Jensen's inequality), the VFE provides an upper bound on the log-evidence – thus, VFE minimization improves both evidence and model fit:

$$\mathcal{F} = \mathop{\mathbb{E}}_{q(\tilde{\boldsymbol{x}})}\left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})}\right] = \mathop{\mathbb{E}}_{q(\tilde{\boldsymbol{x}})}\left[\ln \frac{q(\tilde{\boldsymbol{x}})}{p(\tilde{\boldsymbol{x}}|\tilde{\boldsymbol{y}})}\right] - \ln p(\tilde{\boldsymbol{y}}) \tag{31}$$

Under the previous approximations, the free energy minimization turns into an iterative parameter update through the following expressions:

$$\begin{aligned} \dot{\tilde{\boldsymbol{\mu}}} - \mathcal{D}\tilde{\boldsymbol{\mu}} &= -\partial_{\boldsymbol{\mu}}\mathcal{F} = \partial\tilde{\boldsymbol{g}}^T\tilde{\boldsymbol{\Pi}}_y\tilde{\boldsymbol{\varepsilon}}_y + \partial_{\boldsymbol{\mu}}\tilde{\boldsymbol{f}}^T\tilde{\boldsymbol{\Pi}}_x\tilde{\boldsymbol{\varepsilon}}_x - \mathcal{D}^T\tilde{\boldsymbol{\Pi}}_x\tilde{\boldsymbol{\varepsilon}}_x \\ \dot{\tilde{\boldsymbol{\nu}}} - \mathcal{D}\tilde{\boldsymbol{\nu}} &= -\partial_{\boldsymbol{\nu}}\mathcal{F} = \partial_{\boldsymbol{\nu}}\tilde{\boldsymbol{f}}^T\tilde{\boldsymbol{\Pi}}_x\tilde{\boldsymbol{\varepsilon}}_x - \tilde{\boldsymbol{\Pi}}_v\tilde{\boldsymbol{\varepsilon}}_v \end{aligned} \tag{32}$$

where $\tilde{\boldsymbol{\varepsilon}}_s$, $\tilde{\boldsymbol{\varepsilon}}_x$ and $\tilde{\boldsymbol{\varepsilon}}_v$ are respectively prediction errors of sensations, dynamics and prior:

$$\begin{aligned} \tilde{\boldsymbol{\varepsilon}}_y &= \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{g}}(\tilde{\boldsymbol{\mu}}) \\ \tilde{\boldsymbol{\varepsilon}}_x &= \mathcal{D}\tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\nu}}) \\ \tilde{\boldsymbol{\varepsilon}}_v &= \tilde{\boldsymbol{\nu}} - \boldsymbol{\eta} \end{aligned} \tag{33}$$

Notably, the VFE can also be used to compute motor control signals:

$$\dot{\boldsymbol{a}} = -\partial_{\boldsymbol{a}}\mathcal{F} = -\partial_{\boldsymbol{a}}\tilde{\boldsymbol{g}}^T\tilde{\boldsymbol{\Pi}}_y\tilde{\boldsymbol{\varepsilon}}_y \tag{34}$$

In active inference, a goal is generally encoded as a prior belief: since the generative model specify how the world may evolve, if one has the false belief to be in a particular state, his or her actions will eventually result in looking for those states that make his or her beliefs true. This process allows implementing goal-directed behavior through the VFE, keeping the agent in predictable and safer spaces [17]. Note that this process can be scaled up to construct a hierarchical model, where each level is constantly involved in computing a prediction, comparing it with the state below, and updating its state with the generated prediction error. These simple steps are repeated throughout the whole hierarchy so that, from a prior belief encoding a goal, a cascade of primitive proprioceptive trajectories is generated that is eventually realized by the lowest levels of the motor system.

### 4.2 Planning with the expected free energy

Although active inference in continuous time can deal with real-world problems by keeping track of the instantaneous trajectories that the continuous environment produces, it has several limitations and a narrow use as it cannot easily handle more general types of actions, including decision-making. The minimization of the VFE can only adjust the approximate posterior depending on current – or past – observations, and it does not evaluate future states and outcomes. To endow an agent with this ability, a quantity called *Expected Free Energy* (EFE) is considered [23, 24]. We first augment the generative model with policies $\boldsymbol{\pi}$:

$$p(\boldsymbol{s}, \boldsymbol{o}, \boldsymbol{\pi}) = p(\boldsymbol{o}|\boldsymbol{s}, \boldsymbol{\pi})p(\boldsymbol{s}|\boldsymbol{\pi})p(\boldsymbol{\pi}) \tag{35}$$

where $\boldsymbol{s}$ and $\boldsymbol{o}$ are discrete states and outcomes. Note that the policies are not simple stimulus-response mappings – as in Reinforcement Learning schemes – but sequences of actions. Since action planning involves selecting those policies that lead to the desired priors, we also have to consider future outcomes that have not been observed yet. Thus, the EFE is specifically constructed by conditioning over them, which are therefore treated as hidden states:

$$\mathcal{G}_\pi = \underset{q(\boldsymbol{s},\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} \left[ \ln \frac{q(\boldsymbol{s}|\boldsymbol{\pi})}{p(\boldsymbol{s}, \boldsymbol{o}|\boldsymbol{\pi})} \right] \approx \underset{q(\boldsymbol{s},\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} \left[ \ln \frac{q(\boldsymbol{s})}{q(\boldsymbol{s}|\boldsymbol{o}, \boldsymbol{\pi})} \right] - \underset{q(\boldsymbol{o}|\boldsymbol{\pi})}{\mathbb{E}} [\ln p(\boldsymbol{o}|\boldsymbol{C})] \tag{36}$$

where the probability distribution $p(\boldsymbol{o}|\boldsymbol{C})$ encodes preferred outcomes. The last two terms are respectively called *epistemic* (uncertainty reducing) and *pragmatic* (goal seeking). In practice, this quantity is used by first factorizing the agent's generative model as in POMDPs:

$$p(\boldsymbol{s}_{1:T}, \boldsymbol{o}_{1:T}, \boldsymbol{\pi}) = p(\boldsymbol{s}_1) \cdot p(\boldsymbol{\pi}) \cdot \prod_{\tau=1}^{T} p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau) \cdot \prod_{\tau=2}^{T} p(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1}, \boldsymbol{\pi}) \tag{37}$$

Each of these elements can be represented with categorical distributions:

$$\begin{aligned} p(\boldsymbol{s}_1) &= Cat(\boldsymbol{D}) & p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau) &= Cat(\boldsymbol{A}) \\ p(\boldsymbol{\pi}) &= Cat(\boldsymbol{E}) & p(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1}, \boldsymbol{\pi}) &= Cat(\boldsymbol{B}_{\pi,\tau}) \end{aligned} \tag{38}$$

where $\boldsymbol{D}$ encodes beliefs about the initial state, $\boldsymbol{E}$ encodes the prior over policies, $\boldsymbol{A}$ is the likelihood matrix and $\boldsymbol{B}_{\pi,\tau}$ is the transition matrix. If we assume, under the mean-field approximation, that the approximate posterior factorizes into independent distributions:

$$\begin{aligned} p(\boldsymbol{s}_{1:T}|\boldsymbol{o}_{1:T}, \boldsymbol{\pi}) \approx q(\boldsymbol{s}_{1:T}, \boldsymbol{\pi}) &= q(\boldsymbol{\pi}) \prod_{\tau}^{T} q(\boldsymbol{s}_\tau|\pi) \\ q(\boldsymbol{\pi}) &= Cat(\boldsymbol{\pi}) \\ q(\boldsymbol{s}_\tau|\boldsymbol{\pi}) &= Cat(\boldsymbol{s}_{\pi,\tau}) \end{aligned} \tag{39}$$

we can use the standard technique of variational message passing to infer each posterior $q(\boldsymbol{s}_\tau|\boldsymbol{\pi})$ and combine them into a global posterior $q(\boldsymbol{s}_{1:T}|\boldsymbol{\pi})$. In order to update the posterior about hidden states, we combine the messages from past states, future states, and outcomes, express each term with its sufficient statistics, and finally apply a softmax function to get a proper probability distribution:

$$\boldsymbol{s}_{\pi,\tau} = \sigma(\ln \boldsymbol{B}_{\pi,\tau-1}\boldsymbol{s}_{\pi,\tau-1} + \ln \boldsymbol{B}_{\pi,\tau}^T \boldsymbol{s}_{\pi,\tau+1} + \ln \boldsymbol{A}^T \boldsymbol{o}_\tau) \tag{40}$$

Similarly, in order to update $q(\boldsymbol{\pi})$, we combine the messages from the prior over policies given by the matrix $\boldsymbol{E}$, and from future observations conditioned upon policies; we can approximate the latter by the EFE conditioned at a particular time $\tau$:

$$\begin{aligned} \mathcal{G}_\pi &\approx \sum_\tau D_{KL}[q(\boldsymbol{o}_\tau|\boldsymbol{\pi})||p(\boldsymbol{o}_\tau|\boldsymbol{C})] + \underset{q(\boldsymbol{s}_\tau|\boldsymbol{s}_{\tau-1},\boldsymbol{\pi})}{\mathbb{E}} [H[p(\boldsymbol{o}_\tau|\boldsymbol{s}_\tau)]] \\ &= \sum_\tau \boldsymbol{o}_{\pi,\tau}(\ln \boldsymbol{o}_{\pi,\tau} - \boldsymbol{C}_\tau) + \boldsymbol{s}_{\pi,\tau}\boldsymbol{H} \end{aligned} \tag{41}$$

where:

$$\boldsymbol{o}_{\pi,\tau} = \boldsymbol{A}\boldsymbol{s}_{\pi,\tau} \qquad \boldsymbol{C}_\tau = \ln p(\boldsymbol{o}_\tau|\boldsymbol{C}) \qquad \boldsymbol{H} = -diag(\boldsymbol{A}^T \ln \boldsymbol{A}) \tag{42}$$

Finally, we select the action that is the most likely under all policies:

$$\begin{aligned} \boldsymbol{\pi} &= \sigma(\ln \boldsymbol{E} - \boldsymbol{\mathcal{G}}) \\ u_t &= \arg\max_u \boldsymbol{\pi} \cdot [U_{\pi,t} = u] \end{aligned} \tag{43}$$

### 4.3 Bayesian model reduction in hybrid models

To get a discrete model working with the richness of the continuous input, a form of communication with the continuous framework is required. In particular, since the two levels here operate in different domains, we need a way to obtain a continuous prior from a discrete prediction and, concurrently, to estimate a discrete outcome based on continuous evidence. Both problems can be easily addressed through Bayesian model reduction [37, 36]. Consider a generative model $p(\boldsymbol{\theta}, \boldsymbol{y})$ with parameters $\boldsymbol{\theta}$ and data $\boldsymbol{y}$:

$$p(\boldsymbol{\theta}, \boldsymbol{y}) = p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{44}$$

and an additional distribution $\tilde{p}(\boldsymbol{y}, \boldsymbol{\theta})$, which is a reduced version of the first model if the likelihood of some data is the same under both models and the only difference rests upon the specification of the priors $\tilde{p}(\boldsymbol{\theta})$. We can express the posterior of the reduced model as:

$$\tilde{p}(\boldsymbol{\theta}|\boldsymbol{y}) = p(\boldsymbol{\theta}|\boldsymbol{y})\frac{\tilde{p}(\boldsymbol{\theta})p(\boldsymbol{y})}{p(\boldsymbol{\theta})\tilde{p}(\boldsymbol{y})} \tag{45}$$

The procedure for computing the reduced posterior is the following: first, we integrate over the parameters to obtain the evidence ratio of the two models:

$$\tilde{p}(\boldsymbol{y}) = p(\boldsymbol{y}) \int p(\boldsymbol{\theta}|\boldsymbol{y})\frac{\tilde{p}(\boldsymbol{\theta})}{p(\boldsymbol{\theta})}d\boldsymbol{\theta} \tag{46}$$

Then, we define an approximate posterior $q(\boldsymbol{\theta})$ and we compute the reduced VFE in terms of the full model:

$$\mathcal{F}[\tilde{p}(\boldsymbol{\theta})] \approx \mathcal{F}[p(\boldsymbol{\theta})] + \ln \mathbb{E}_q \left[\frac{\tilde{p}(\boldsymbol{\theta})}{p(\boldsymbol{\theta})}\right] \tag{47}$$

Correspondingly, the approximate posterior of the reduced model can be written in terms of the full model:

$$\ln \tilde{q}(\boldsymbol{\theta}) = \ln q(\boldsymbol{\theta}) + \ln \frac{\tilde{p}(\boldsymbol{\theta})}{p(\boldsymbol{\theta})} - \ln \mathbb{E}_q \left[\frac{\tilde{p}(\boldsymbol{\theta})}{p(\boldsymbol{\theta})}\right] \tag{48}$$

In order to apply this technique to the communication between discrete and continuous models, we first consider as parameters the hidden causes $\boldsymbol{v}$ and denote with $\boldsymbol{y}$ their continuous observations. We then assume that the full prior probability depends on a discrete outcome:

$$p(\boldsymbol{v}|\boldsymbol{o}) = \mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Pi}_v^{-1}) \tag{49}$$

In this way, the parameter $\boldsymbol{\eta}$ acts as a prior over the hidden causes. We also assume that the agent maintains $M$ reduced models with prior:

$$\tilde{p}(\boldsymbol{v}|\boldsymbol{o}, m) = \mathcal{N}(\boldsymbol{\eta}_m, \boldsymbol{\Pi}_m^{-1}) \tag{50}$$

These models correspond to specific hypotheses or assumptions that the agent believes to be possible causes of the sensorium. In other words, it is the reduced prior $\boldsymbol{\eta}_m$ that encodes a prior configuration for the continuous model, and each of them is associated with a particular discrete outcome $o_{\tau,m}$. The latter are computed by marginalizing $\boldsymbol{o}_{\pi,\tau}$ over all policies:

$$\boldsymbol{o}_\tau = \sum_\pi \pi_\pi \boldsymbol{o}_{\pi,\tau} \tag{51}$$

This means that the prior of the full model, which represents the actual agent's parameters of the current state of the world, is obtained by simply performing a Bayesian model average between every specified model:

$$\boldsymbol{\eta} = \sum_m o_{\tau,m}\boldsymbol{\eta}_m \tag{52}$$

thus transforming discrete probabilities into a continuous value. This quantity will then bias the belief over hidden causes $\boldsymbol{\nu}$ as in Equations 32-33. Similarly, having defined the full and reduced approximate posteriors:

$$\begin{aligned} q(\boldsymbol{v}) &= \mathcal{N}(\boldsymbol{\nu}, \boldsymbol{P}_v^{-1}) \\ \tilde{q}(\boldsymbol{v}|m) &= \mathcal{N}(\boldsymbol{\nu}_m, \boldsymbol{P}_m^{-1}) \end{aligned} \tag{53}$$

the ascending message that will act as a discrete observation for the higher-level model is computed through Bayesian model comparison, i.e., by comparing the probability that each discrete outcome model can be the true cause of the sensory input. In this case, the free energy computed in Equation 47 acts as a hint to how well a reduced model explains continuous observations, and it corresponds to the prior surprise plus the log-evidence for each outcome model accumulated over time:

$$\mathcal{F}_m = -\ln o_{\tau,m} - \int_0^T \mathcal{L}_m dt \tag{54}$$

The Laplace approximation [57] leads to a very simple form of the second RHS term:

$$\mathcal{L}_m = \frac{1}{2}\ln|\boldsymbol{\Pi}_m \boldsymbol{P}_v \boldsymbol{P}_m^{-1}\boldsymbol{\Sigma}_v| - \frac{1}{2}(\boldsymbol{\nu}^T \boldsymbol{P}_v \boldsymbol{\nu} - \boldsymbol{\nu}_m^T \boldsymbol{P}_m \boldsymbol{\nu}_m - \boldsymbol{\eta}^T \boldsymbol{\Pi}_v \boldsymbol{\eta} + \boldsymbol{\eta}_m^T \boldsymbol{\Pi}_m \boldsymbol{\eta}_m) \tag{55}$$

where the mean and precision of the $m$th reduced model are:

$$\begin{aligned} \boldsymbol{\nu}_m &= \boldsymbol{P}_m^{-1}(\boldsymbol{P}_v \boldsymbol{\nu} - \boldsymbol{\Pi}_v \boldsymbol{\eta} + \boldsymbol{\Pi}_m \boldsymbol{\eta}_m) \\ \boldsymbol{P}_m &= \boldsymbol{P}_v - \boldsymbol{\Pi}_v + \boldsymbol{\Pi}_m \end{aligned} \tag{56}$$

After having computed the free energy associated with each model, they are normalized through a softmax function to get a proper probability, which is finally used to estimate the current discrete state:

$$\begin{aligned} \boldsymbol{r}_\tau &= \sigma(-\boldsymbol{\mathcal{F}}) \\ \boldsymbol{s}_{\pi,\tau} &= \sigma(\ln \boldsymbol{B}_{\pi,\tau-1}\boldsymbol{s}_{\pi,\tau-1} + \ln \boldsymbol{B}_{\pi,\tau}^T \boldsymbol{s}_{\pi,\tau+1} + \ln \boldsymbol{A}^T \boldsymbol{r}_\tau) \end{aligned} \tag{57}$$

Each value of $\boldsymbol{r}_\tau$ then indicates how probable is a discrete outcome based on the accumulated continuous evidence. Contrary to the discrete-only case, in a hybrid model the discrete actions are only implicitly used to compute the posterior over policies, and the link between the two frameworks is possible by associating each discrete outcome model $o_{\tau,m}$ to a particular continuous prior $\boldsymbol{\eta}_m$.

## Acknowledgments

This research received funding from the European Union's Horizon H2020-EIC-FETPROACT-2019 Programme for Research and Innovation under Grant Agreement 951910 to I.S. and the Italian Ministry for Research MIUR under Grant Agreement PRIN 2017KZNZLN to I.S. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References

[1] Thomas Parr, Rajeev Vijay Rikhye, Michael M. Halassa, and Karl J. Friston. Prefrontal Computation as Active Inference. *Cerebral Cortex*, 30(2):682–695, 2020.

[2] Karl J. Friston, Richard Rosch, Thomas Parr, Cathy Price, and Howard Bowman. Deep temporal models and active inference. *Neuroscience and Biobehavioral Reviews*, 77(November 2016):388–402, 2017.

[3] Thomas Brashers-Krug, Reza Shadmehr, and Emilio Bizzi. Consolidation in human motor memory. *Nature*, 382(5505):252–255, 1996.

[4] Daniel E Callan and Eiichi Naito. from Expert and Novice Athletes. 27(4):183–188, 2014.

[5] Francesco Di Russo, Sabrina Pitzalis, Teresa Aprile, and Donatella Spinelli. Effect of practice on brain activity: An investigation in top-level rifle shooters. *Medicine and Science in Sports and Exercise*, 37(9):1586–1593, 2005.

[6] F. Fattapposta, G. Amabile, M. V. Cordischi, D. Di Venanzio, A. Foti, F. Pierelli, C. D'Alessio, F. Pigozzi, A. Parisi, and C. Morrocutti. Long-term practice effects on a new skilled motor learning: An electrophysiological study. *Electroencephalography and Clinical Neurophysiology*, 99(6):495–507, 1996.

[7] Lucio Marinelli, Angelo Quartarone, Mark Hallett, Giuseppe Frazzitta, and Maria Felice Ghilardi. The many facets of motor learning and their relevance for Parkinson's disease. *Clinical Neurophysiology*, 128(7):1127–1141, 2017.

[8] Elisenda Bueichekú, Anna Miró-Padilla, María Ángeles Palomar-García, Noelia Ventura-Campos, María Antonia Parcet, Alfonso Barrós-Loscertales, and César Ávila. Reduced posterior parietal cortex activation after training on a visual search task. *NeuroImage*, 135:204–213, 2016.

[9] J. R. Wolpaw. Spinal cord plasticity in acquisition and maintenance of motor skills. *Acta Physiologica*, 189(2):155–169, 2007.

[10] Alice Nieuwboer, Lynn Rochester, Liesbeth Müncks, and Stephan P. Swinnen. Motor learning in Parkinson's disease: limitations and potential for rehabilitation. *Parkinsonism and Related Disorders*, 15(SUPPL. 3):53–58, 2009.

[11] Giovanni Pezzulo, Francesco Rigoli, and Karl J. Friston. Hierarchical Active Inference: A Theory of Motivated Control. *Trends in Cognitive Sciences*, 22(4):294–306, 2018.

[12] Ann M. Graybiel. Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31:359–387, 2008.

[13] R. J. Nudo, G. W. Milliken, W. M. Jenkins, and M. M. Merzenich. Use-Dependent Alterations of Movement Representations in Primary Motor Cortex of Adult Squirrel Monkeys. *The Journal of Neuroscience*, 16(2):785–807, 1996.

[14] Reza Shadmehr and Henry H. Holcomb. Neural correlates of motor memory consolidation. *Science*, 277(5327):821–825, 1997.

[15] D M Wolpert, Z Ghahramani, and M Jordan. An internal model for sensorimotor integration - Wolpert et al. (1995).pdf. *Science*, 269(5):1880–1882, 1995.

[16] Karl Friston and Stefan Kiebel. Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521):1211–1221, 2009.

[17] Karl J. Friston, Jean Daunizeau, James Kilner, and Stefan J. Kiebel. Action and behavior: A free-energy formulation. *Biological Cybernetics*, 102(3):227–260, 2010.

[18] Thomas Parr, Giovanni Pezzulo, and Karl J Friston. *Active inference: the free energy principle in mind, brain, and behavior*. Cambridge, MA: MIT Press, 2021.

[19] Rajesh P.N. Rao and Dana H. Ballard. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87, 1999.

[20] Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[21] Matteo Priorelli, Federico Maggiore, Antonella Maselli, Francesco Donnarumma, Domenico Maisto, Francesco Mannella, Ivilin Peev Stoianov, and Giovanni Pezzulo. Modeling motor control in continuous-time Active Inference: a survey. *IEEE Transactions on Cognitive and Developmental Systems*, pages 1–15, 2023.

[22] Rick A. Adams, Stewart Shipp, and Karl J. Friston. Predictions not commands: Active inference in the motor system. *Brain Structure and Function*, 218(3):611–643, 2013.

[23] Lancelot Da Costa, Thomas Parr, Noor Sajid, Sebastijan Veselic, Victorita Neacsu, and Karl Friston. Active inference on discrete state-spaces: A synthesis. *Journal of Mathematical Psychology*, 99, 2020.

[24] Ryan Smith, Karl J. Friston, and Christopher J. Whyte. A step-by-step tutorial on active inference and its application to empirical data. *Journal of Mathematical Psychology*, 107:102632, 2022.

[25] Raphael Kaplan and Karl J. Friston. Planning and navigation as active inference. *Biological Cybernetics*, 112(4):323–343, 2018.

[26] Planning as inference. *Trends in Cognitive Sciences*, 16(10):485–488, 2012.

[27] Marc Toussaint. Probabilistic inference as a model of planned behavior. *Künstliche Intelligenz*, 3/09:23–29, 2009.

[28] Sherin Grimbergen, S S Grimbergen, · C Van Hoof, · P Mohajerin Esfahani, and · M Wisse. Active Inference for State Space Models: A Tutorial. 2019.

[29] Noor Sajid, Philip J. Ball, Thomas Parr, and Karl J. Friston. Active inference: demystified and compared. *Neural Computation*, 33(3):674–712, 2021.

[30] Philipp Schwartenbeck, Thomas H.B. Fitzgerald, Christoph Mathys, Ray Dolan, Martin Kronbichler, and Karl Friston. Evidence for surprise minimization over value maximization in choice behavior. *Scientific Reports*, 5:1–14, 2015.

[31] Philipp Schwartenbeck, Johannes Passecker, Tobias U. Hauser, Thomas H.B. Fitzgerald, Martin Kronbichler, and Karl J. Friston. Computational mechanisms of curiosity and goal-directed exploration. *eLife*, 8:1–45, 2019.

[32] Karl J. Friston, Jean Daunizeau, and Stefan J. Kiebel. Reinforcement learning or active inference? *PLoS ONE*, 4(7), 2009.

[33] Karl J Friston, S Samothrakis, and Read Montague. Active inference and agency: optimal control without cost functions. *Biological cybernetics*, (106):523–541, 2012.

[34] Karl J. Friston, Thomas Parr, and Bert de Vries. The graphical brain: Belief propagation and active inference. 1(4):381–414, 2017.

[35] Thomas Parr and Karl J. Friston. The Discrete and Continuous Brain: From Decisions to Movement—And Back Again. (September):2319–2347, 2018.

[36] Karl Friston, Thomas Parr, and Peter Zeidman. Bayesian model reduction. pages 1–32, 2018.

[37] Karl Friston and Will Penny. Post hoc Bayesian model selection. *NeuroImage*, 56(4):2089–2099, 2011.

[38] T. Parr and K. J. Friston. The computational pharmacology of oculomotion. *Psychopharmacology (Berl.)*, 236(8), August 2019.

[39] Alexander Tschantz, Laura Barca, Domenico Maisto, Christopher L. Buckley, Anil K. Seth, and Giovanni Pezzulo. Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using active inference. *Biological Psychology*, 169, 2022.

[40] K. J. Friston, N. Sajid, D. R. Quiroga-Martinez, T. Parr, C. J. Price, and E. Holmes. Active listening. *Hear. Res.*, 399:107998, January 2021.

[41] Stefano Ferraro, Toon Van de Maele, Pietro Mazzaglia, Tim Verbelen, and Bart Dhoedt. Disentangling shape and pose for object-centric deep active inference models, 2022.

[42] Matteo Priorelli and Ivilin Peev Stoianov. Flexible Intentions: An Active Inference Theory. *Frontiers in Computational Neuroscience*, 17:1 – 41, 2023.

[43] Matteo Priorelli, Giovanni Pezzulo, and Ivilin Peev Stoianov. Deep kinematic inference affords efficient and scalable control of bodily movements. *Proceedings of the National Academy of Sciences*, 120, 2023.

[44] Matteo Priorelli and Ivilin Peev Stoianov. Efficient motor learning through action-perception cycles in deep kinematic inference. In *Active Inference*, pages 59–70. Springer Nature Switzerland, 2024.

[45] M. Priorelli and I.P. Stoianov. Dynamic inference by model reduction. *bioRxiv*, 2023.

[46] Thomas H.B. FitzGerald, Rosalyn J. Moran, Karl J. Friston, and Raymond J. Dolan. Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. *NeuroImage*, 107:219–228, 2015.

[47] Janelle Weaver. Motor Learning Unfolds over Different Timescales in Distinct Neural Systems. *PLoS Biology*, 13(12):1–2, 2015.

[48] Matteo Priorelli and Ivilin Peev Stoianov. Intention Modulation for Multi-Step Tasks in Continuous Time Active Inference. In *Active Inference, Third International Workshop, IWAI 2022, Grenoble, France, Sept 19, 2022*, 2022.

[49] Emanuel Todorov. Optimality principles in sensorimotor control. *Nature Neuroscience*, 7:907–915, 2004.

[50] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[51] Jakob Hohwy. *The Predictive Mind*. Oxford University Press UK, 2013.

[52] Daniel Baldauf, He Cui, and Richard A. Andersen. The posterior parietal cortex encodes in parallel both goals for double-reach sequences. *Journal of Neuroscience*, 28(40):10081–10089, 2008.

[53] Stefan Schaal, Dagmar Sternad, Rieko Osu, and Mitsuo Kawato. Rhythmic arm movement is not discrete. *Nature Neuroscience*, 7(10):1137–1144, 2004.

[54] A. Maselli, P. Lanillos, and G. Pezzulo. Active inference unifies intentional and conflict-resolution imperatives of motor control. *PLOS Comput. Biol*, 18(6), 2022.

[55] Francesco Mannella, Federico Maggiore, Manuel Baltieri, and Giovanni Pezzulo. Active inference through whiskers. *Neural Networks*, 144:428–437, 2021.

[56] Marshall G. Shuler and Mark F. Bear. Reward Timing in the Primary Visual Cortex. *Science*, (March):1606–1610, 2015.

[57] Karl Friston, Jérémie Mattout, Nelson Trujillo-Barreto, John Ashburner, and Will Penny. Variational free energy and the Laplace approximation. *NeuroImage*, 34(1):220–234, 2007.

# 5 Supplementary Materials

## 5.1 Simulating affordances

Defining continuous intentions also in the intrinsic domain has the advantage that high hierarchical levels can impose priors not only on hand positions, but also on specific configurations of the arm and fingers. This is useful when the

agent has to grasp an object in a particular manner depending on the task considered (e.g., power grip vs precision grip). Figure 7 shows a simple demonstration of this behavior: here, the agent is required to grasp the object with a counterclockwise wrist rotation, and then place it at the goal position with a clockwise rotation. The arm starts with the fingers closed and the object belief is initialized at the hand position. Then it starts reaching the object while slowly opening the hand and preparing the correct wrist rotation. Being subject to multi-domain constraints, this configuration is maintained until the moving object is grasped. Note that there is a tradeoff between reaching the object and maintaining the imposed constraint on the wrist, as can be seen in the middle of the movement where the agent waits until the object can be grasped with the correct rotation even if the latter is already in its peripersonal space. When the object is grasped, it rotates again the wrist by 180 degrees clockwise and correctly places the object in the goal position.

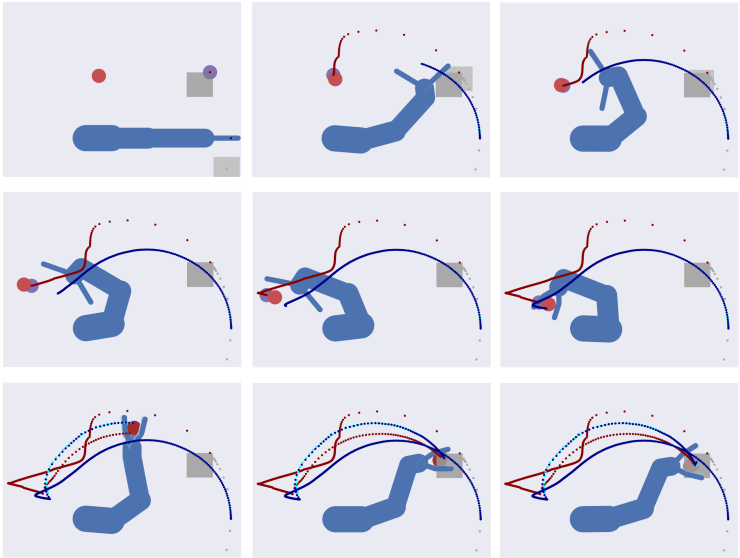


Figure 7: Sequence of time frames of the grasping task with affordances. Real and estimated object positions are represented by red and purple circles, real and estimated goal positions by grey and dark grey squares, and real and estimated limb positions in blue and cyan. Hand, object, and goal belief trajectories are represented respectively with a blue, red, and grey dotted line.

### 5.2 Supplementary Movie 1

Simulation of the dynamic pick-and-place task with a hybrid model.

### 5.3 Supplementary Movie 2

Simulation of the dynamic pick-and-place task with a continuous-only model. Also see Figure 4.

### 5.4 Supplementary Movie 3

Simulation of the dynamic pick-and-place task with affordances. Also see Figure 7.