

IG3A: An Information-Gain-Based Active Authority Allocation Framework for Shared Control Under Human Bounded Rationality

Yun-Sheng Zhao¹, Yun-Bo Zhao¹, *Senior Member, IEEE*, Pengfei Li², Cegang Li, Shuyue Jiang¹, Kangcheng Wang¹, Shumi Zhao¹, and Yu Kang¹, *Senior Member, IEEE*

Abstract—Shared control aims to enhance human–machine collaboration by dynamically integrating complementary human and machine intelligence. However, within such frameworks, a fundamental challenge lies in dealing with the uncertainty regarding human intention arising from human bounded rationality. Although prior studies have made substantial progress in representing and using intention uncertainty, the effect of the current sharing decision on future intention uncertainty has received less explicit attention. To address this, we propose the Information-Gain-based Active Authority Allocation (IG3A) framework, formulated within a linear blending setting, to establish a principled basis for authority arbitration. The core innovation lies in constructing a tractable surrogate objective that explicitly balances task exploitation and intention exploration, thereby avoiding the computational intractability of long-horizon belief-space optimization and transforming authority allocation into an active optimization process. Enabled by a Monte Carlo-based online solution scheme, IG3A evaluates candidate authority levels in real time and selects the authority level according to the task-exploration trade-off. Simulated-human LunarLander experiments, together with human-in-the-loop LunarLander and physical robotic-arm experiments, show that IG3A can improve overall shared-control performance, particularly under uncertain or changing intention evidence, by explicitly trading off immediate task execution and active intention exploration.

Overall, the proposed framework offers a principled basis for information-aware authority allocation in shared control under human bounded rationality, with potential for extension to more general settings.

Note to Practitioners—This paper presents a shared control framework for teleoperated robotic systems, such as remote manipulators in hazardous environments and assistive devices such as smart wheelchairs. In these applications, operator commands may be imprecise, inconsistent, or delayed due to physical fatigue, cognitive limitations, limited task visibility, or unfamiliar system dynamics. The proposed approach serves as an intelligent arbitration unit that can be integrated into existing continuous teleoperation loops. The controller adjusts the authority level between the human operator and the autonomous agent by considering both immediate task execution and the expected reduction of intention uncertainty. For implementation, the system requires a perception module that provides a finite set of candidate targets or intentions, an autonomy module that can generate goal-conditioned assistive actions, and a continuous input interface such as a 6-DoF joystick or SpaceMouse. The practical value of the framework lies in providing an explicit, interpretable, and deployable authority-arbitration mechanism that can improve overall shared-control performance by trading off immediate task execution and active intention exploration under uncertain or changing user intentions. The present implementation assumes a predefined finite intention set and does not automatically handle open-set targets that are absent from the perception module.

Index Terms—Shared control, bounded rationality, information gain, authority allocation.

Received 8 January 2026; revised 19 May 2026 and 12 July 2026; accepted 2 September 2026. Date of publication 9 September 2026; date of current version 15 September 2026. This article was recommended for publication by Associate Editor A.-T. Nguyen and Editor A. Pietrabissa upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 62473352 and in part by Anhui Provincial Natural Science Foundation under Grant 2208085UD09. (*Corresponding author: Yun-Bo Zhao.*)

Yun-Sheng Zhao, Pengfei Li, Cegang Li, and Shuyue Jiang are with the Department of Automation, University of Science and Technology of China, Hefei 230026, China (e-mail: zys1030@mail.ustc.edu.cn; puffylee@ustc.edu.cn; lcg123@mail.ustc.edu.cn; jiangsuy@mail.ustc.edu.cn).

Yun-Bo Zhao is with the Department of Automation, University of Science and Technology of China, Hefei 230026, China, also with the Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China, and also with the Institute of Advanced Technology, University of Science and Technology of China, Hefei 230088, China (e-mail: ybzhao@ustc.edu.cn).

Kangcheng Wang and Shumi Zhao are with the Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China (e-mail: kcwang@ustc.edu.cn; shumi_zhao@163.com).

Yu Kang is with the Department of Automation and Anhui Province Key Laboratory of Intelligent Low-Carbon Information Technology and Equipment, University of Science and Technology of China, Hefei 230026, China, and also with the School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China (e-mail: kangduyu@ustc.edu.cn).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TASE.2026.3732136>, provided by the authors.

Digital Object Identifier 10.1109/TASE.2026.3732136

I. INTRODUCTION

SHARED control serves as a cutting-edge paradigm for human–machine collaboration, characterized by the dynamic integration of human and machine intelligence [1], [2], [3]. Its core rationale lies in synergizing complementary strengths: the paradigm combines the superior global perception and experience-based decision-making of humans with the powerful capabilities in precise computation and rapid execution of machines. Through this synergy, shared control can be leveraged to achieve diverse objectives, ranging from mitigating operator workload (spanning both cognitive and physical dimensions) to enhancing task safety and execution performance [4], [5]. Given these significant benefits, the paradigm has shown great potential across a wide range of domains, such as advanced driver-assistance systems (ADAS) [6], [7], medical surgery [8], and robotic manipulators [9].

A fundamental challenge in shared control frameworks, particularly those designed to optimize task performance, lies in dealing with the uncertainty regarding human intention arising from human bounded rationality. Simon introduced the concept of *bounded rationality* to characterize systematic deviations from perfect rationality caused by limitations in both available information (sensing) and computational capacity (decision-making) [10]. While perceptual constraints contribute to these deviations, this work specifically addresses the decision-making limitations: even when task information is accessible, human operators struggle to comprehend complex system dynamics and plan long-horizon trajectories in real time, forcing them to rely on simplified heuristics rather than global optimization [11]. Under an assumption of perfect rationality, human inputs would strictly follow the optimal policy associated with the user's intended objective, making them highly informative evidence of the underlying intention. However, bounded decision-making introduces suboptimality and variability that obscure this mapping: control signals no longer reliably correspond to the optimal action of a single objective [12], [13]. Consequently, the autonomous system is forced to operate under inherent uncertainty, as the human's true intention becomes partially unobservable due to the degradation in decision quality.

Although prior studies have made substantial progress in representing and using intention uncertainty in shared control, the information value of the current sharing decision itself has received less explicit attention. Existing studies include predict-then-act assistance [14], [15], belief-space assistance and information-aware decision making [16], [17], and learning-based or adaptive assistance [18]. These lines of work show that intention uncertainty can be estimated, represented, and used for assistance generation, belief-space planning, active disambiguation, or learned intervention. However, in continuous shared-control interfaces where human and autonomous commands coexist, the current sharing decision is not merely a passive response to the current belief; it also affects the executed system action, the subsequent state, and the future human input from which intention is inferred. Therefore, if arbitration focuses only on the current belief or immediate task assistance, it may prematurely reinforce an incorrect intention estimate, slow down subsequent belief convergence, and increase human-machine conflict when the user's intention is ambiguous or changing. This raises a more specific question: how should the current sharing decision itself use the existing belief to balance immediate task execution with the expected reduction of future intention uncertainty?

To address this coupling between authority allocation and future intention uncertainty, this paper proposes the Information-Gain-based Active Authority Allocation (IG3A) framework, which explicitly incorporates the expected reduction of intention uncertainty into the authority-allocation objective. At a conceptual level, our central premise is that, under human bounded rationality, explicitly encouraging active intention exploration is beneficial for shared control, because reducing intention uncertainty can improve downstream assistance quality. This idea is not limited to

a specific arbitration structure, but reflects a broader design principle for shared control under uncertainty. In this work, this idea is developed within a linear blending setting, which provides a simple and interpretable execution interface for shared control. Structurally, this framework formulates shared control under bounded rationality between the human operator and the autonomous agent—defined herein as any computational system sharing control authority [19], [20]—as a joint optimization problem characterized by information-decision asymmetry. The core innovation lies in constructing a tractable surrogate objective that explicitly balances task exploitation and intention exploration, thereby avoiding the computational intractability of the long-horizon belief-space optimization. Crucially, this implies that the authority allocation strategy is determined by the optimization of this surrogate, rather than relying only on manually designed heuristic rules.

The main contributions of this article are summarized as follows:

- A systematic information-aware authority-allocation framework, IG3A, is proposed for linear-blending shared control under human bounded rationality, providing a principled formulation for optimizing the authority allocation beyond purely rule-based mappings. The core innovation lies in formulating authority allocation as a joint optimization problem characterized by information-decision asymmetry between the human and the autonomous agent.
- A tractable surrogate objective is constructed for the otherwise intractable long-horizon belief-space optimization problem, enabling authority allocation to improve overall task performance through an explicit trade-off between immediate task execution and the expected reduction of future intention uncertainty. The key innovation lies in incorporating an information gain term into the objective function, thereby establishing a computable mechanism that uses the authority coefficient not only for control arbitration but also for active intention exploration.
- A computationally efficient online solution algorithm is devised, bridging the gap between the proposed information-aware formulation and practical real-time deployment. By evaluating candidate authority levels with Monte Carlo approximation, this approach makes the surrogate optimization executable in real time and enables online selection of the authority level according to the task-exploration trade-off.

II. RELATED WORK

This section reviews related work to clarify the position of IG3A within shared control research. Although *shared control* and *shared autonomy* are sometimes used with different emphases, both terms are often used in the literature to describe human-machine systems in which human input and autonomous assistance jointly affect task execution. For consistency, this paper mainly uses *shared control*, while also discussing closely related shared-autonomy studies.

The review follows the problem logic of this paper: bounded rationality motivates a non-ideal view of human commands; imperfect human commands lead to intention

uncertainty in shared control; and IG3A further focuses on the information–decision coupling in authority allocation, where the current control-sharing decision affects both immediate task performance and future reduction of intention uncertainty.

A. Bounded Rationality in Shared Control

Human operators in shared control are often goal-directed, but their commands are not necessarily optimal. Recent shared-control and shared-autonomy studies have increasingly recognized this issue under the notions of bounded, suboptimal, or imperfect human rationality. Some works explicitly incorporate bounded rationality into shared-control modeling, such as differential-game-based shared control with incomplete information [21], safe adaptive dynamic programming for bounded-rational human–machine stabilization [13], and POMDP-based shared autonomy with hidden goal and rationality states for mitigating over-assistance [22]. Related work on mutual adaptation also treats the human as an adaptive but imperfect partner rather than a perfectly rational controller [12]. These studies show that bounded rationality is not merely a behavioral-science background assumption, but an important modeling factor in shared-control systems.

A broader group of studies does not explicitly use the term bounded rationality, but still accounts for non-ideal human behavior through skill, ability, workload, fatigue, trust, attentiveness, or other human-factor models. For example, shared-driving, teleoperation, and assistive-control methods adapt assistance or authority according to time-varying driver characteristics, operator skill, stochastic skill levels, workload, fatigue, mutual trust, abnormal driver states, or attentiveness [23], [24], [25], [26], [27]. In much of this literature, human imperfections are modeled, estimated, or adapted to, and controllers are designed to remain robust to such imperfections, with the aim of improving stability, safety, assistance adaptation, conflict mitigation, task performance, or user experience.

IG3A focuses on a complementary implication of bounded rationality. In task-oriented shared control, bounded-rational human commands are simultaneously control inputs and imperfect evidence of the user’s underlying intention. Thus, the current control-sharing decision affects not only the immediate executed action, but also the subsequent evolution of intention belief. This motivates the use of expected information gain as an intermediate decision value, enabling the arbitrator to balance immediate task assistance with the expected benefit of reducing future intention uncertainty.

B. Shared Control Under Intention Uncertainty

Beyond works that explicitly model bounded rationality, a large body of shared-control and shared-autonomy research addresses intention uncertainty caused by imperfect or ambiguous human inputs. This motivates probabilistic intention inference and uncertainty-aware assistance in task-oriented shared control.

1) *Predict-Then-Act Methods*: A common strategy for handling intention uncertainty in shared control and shared autonomy is to first infer the user’s intention, goal, or task state, and then use this estimate to generate assistance, select

a skill, plan a trajectory, or adjust the autonomy level. This predict-then-act structure is modular and interpretable, because the intention-inference module and the downstream assistance module can be designed and evaluated separately. Classical assistive teleoperation and policy-blending methods follow this general idea by estimating the user’s likely goal and providing goal-directed assistance [28]. Probabilistic intention-recognition methods further formalize this stage by recursively maintaining a distribution over candidate goals from observed user inputs [14]. Recent work has extended this modular structure to more open-ended settings, such as zero-shot intention recognition and vision-language-model-based assistive teleoperation [29], [30].

However, the inferred intention or confidence is usually used as an input to a downstream controller, planner, or arbitration rule, rather than as part of an objective that evaluates how the current control-sharing decision affects future belief evolution.

2) *Belief-Space Assistance and Information-Aware Decision Making*: Another line of work uses intention uncertainty directly in the shared-autonomy decision process. POMDP-style shared autonomy maintains a belief over possible user goals and selects assistance by reasoning over this belief, rather than committing to a single predicted intention. Hind-sight optimization is a representative approximation that evaluates assistance under multiple possible goals and combines the resulting values according to the current belief [16]. Recent POMDP-based shared-autonomy work has further applied this idea to robotic manipulation under human intention uncertainty [9]. These methods provide a principled way to account for goal uncertainty in assistance generation. However, their decision variables are usually assistive actions, policies, or trajectories, and their practical objectives mainly focus on expected task performance under the current or predicted belief.

More explicitly information-aware methods use robot actions, control modes, or interaction contexts to make future user inputs more informative. Brooks and Szafir [17] balance goal-oriented actions with information-gathering actions by moving the robot toward states that are expected to reduce uncertainty about the user’s goal. Gopinath and Argall [31] study intention disambiguation through mode or interface-level assistance, and subsequent work further formulates information-theoretic disambiguation through contextual nudges in a turn-taking shared-control protocol [32]. These studies demonstrate that uncertainty reduction can be useful for shared autonomy. However, they typically realize this idea through informative actions, modes, contexts, or autonomous turns, rather than through the online control-sharing decision when human and autonomous commands coexist.

3) *Learning-Based and Adaptive Assistance*: Recent shared-autonomy methods increasingly use data-driven techniques to learn assistive policies, intervention rules, latent action representations, or personalized control strategies. Deep reinforcement learning has been used to learn shared-autonomy policies directly from interaction data [33], while latent-action and repeated-interaction methods improve adaptability by learning compact action representations or personalized assistance across tasks [19], [34]. More recent

methods use diffusion-based assistance, selective intervention, Bayesian optimization, or policy fine-tuning to adapt shared-control behavior to task conditions, user preferences, or safety requirements [18], [35], [36], [37]. These methods provide strong flexibility and represent an important recent trend in shared autonomy, especially when the assistance policy or intervention strategy can be learned from demonstrations, interaction data, or task-specific training.

This line of work is complementary to IG3A rather than directly equivalent to it. Learning-based methods typically operate at the level of action generation, intervention, policy learning, or long-term adaptation, whereas IG3A focuses on explicit belief-based online arbitration.

C. Positioning of the Proposed Framework

The reviewed studies show that bounded-rational human behavior, intention uncertainty, belief-space assistance, information-aware disambiguation, and learning-based adaptation have been addressed from different perspectives in shared control. These works differ along several dimensions, including their theoretical framework, how human uncertainty is represented, what decision variable is optimized, and what optimization objective is adopted. For example, existing methods may be formulated as modular intention-recognition-and-assistance pipelines, belief-space or POMDP-style decision processes, information-aware disambiguation mechanisms, or data-driven policy-learning frameworks. Uncertainty about the human may be represented as an inferred goal, a goal belief, a cognitive or skill-related state, or an implicit learned representation, while the optimized decision may be an assistive action, mode, trajectory, intervention rule, policy, or adaptation parameter. Their objectives may emphasize task performance, safety, stability, usability, uncertainty reduction, or long-term adaptation. IG3A occupies a different point in this space by treating control sharing as an information-aware arbitration problem rather than only as task execution, assistance selection, or policy learning.

This paper instantiates this perspective in a linear blending interface, where the authority coefficient provides a simple and interpretable decision variable. Within this setting, IG3A introduces expected information gain as an intermediate decision value, so that authority allocation can balance immediate task assistance with the expected benefit of reducing future intention uncertainty. Thus, the proposed framework complements POMDP-style planning, active-disambiguation methods, and learning-based shared autonomy by providing an explicit belief-based arbitration mechanism for shared control under bounded-rational human inputs.

III. PROBLEM FORMULATION

This section formalizes the shared control problem under bounded rationality to mathematically characterize the core challenge identified in the Introduction: the uncertainty in intention and execution sub-optimality arising from bounded rationality. By defining the system dynamics and objective functions, we aim to isolate the specific impact of human

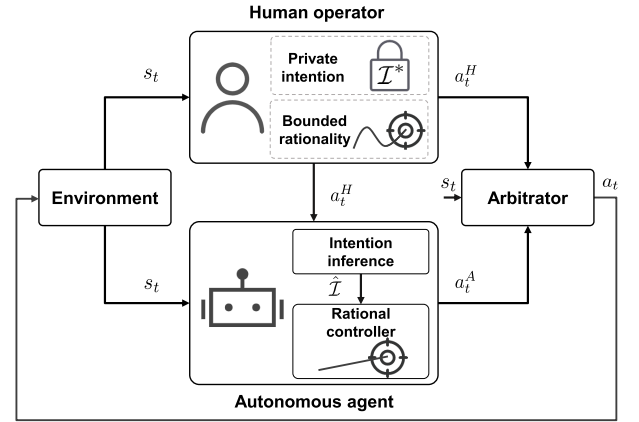


Fig. 1. Shared-control formulation under bounded rationality. The human acts according to a private intention and bounded-rational behavior, whereas the agent maintains an inferred belief and generates assistive control. The arbitrator fuses human and autonomous inputs to produce the final action.

bounded rationality on control efficacy. This formalism serves not merely to establish notation, but to elucidate the limitations of heuristic arbitration in addressing the coupled uncertainty in this setting, thereby motivating the optimization-based framework proposed in Section IV.

A. Mathematical Modeling of the System

As illustrated in Fig. 1, we consider a standard shared control architecture [28] and explicitly characterize the information-decision asymmetry between the human and the agent. The dynamic interaction proceeds as follows: At each time step t , based on the current environmental state s_t , the **human operator** generates a control input a_t^H driven by a private intention $\mathcal{T}^*$. Upon observing a_t^H , the **autonomous agent** utilizes its internal **intention inference** module to update the belief b_t regarding $\mathcal{T}^*$, and subsequently generates an optimal assistive action a_t^A via its **controller**. The inputs from the human and the agent are fused by the **arbitrator** to produce the final system action a_t . While arbitration can generally be modeled as a mapping $a_t = \Phi(a_t^H, a_t^A)$ [16], [38], in this work, we adopt the **linear blending** paradigm [28], [39] for its computational efficiency:

$$a_t = \lambda_t \cdot a_t^H + (1 - \lambda_t) a_t^A, \quad (1)$$

where $\lambda_t \in [0, 1]$ denotes the authority coefficient assigned to the human input, with larger values indicating greater human authority. This linear superposition implicitly requires that the human and the agent operate in a shared action space $\mathcal{A} \subseteq \mathbb{R}^m$ (e.g., end-effector velocity). Finally, the executed action a_t drives the environment dynamics: $s_{t+1} \sim p(s_{t+1} | s_t, a_t)$.

Within this architecture, the decision-making mechanisms of the two agents exhibit an information-decision asymmetry, which is the core setting of this paper, modeled as follows:

- **Human Operator (Full Information, Bounded Rationality):** The human operator is modeled as a goal-directed agent possessing a private intention $\mathcal{T}^* \in \Theta$. This decision-making process is functionally decomposed into two stages:

1) Ideal Objective: Following common formulations in inverse reinforcement learning [40] and shared control [15], the operator aims to maximize a long-term cumulative reward:

$$J(\mathcal{I}^*) = \mathbb{E} \left[\sum_{k=0}^T \gamma^k R_k(s_k, a_k; \mathcal{I}^*) \mid \mathcal{I}^* \right], \quad (2)$$

where R_k , γ , and T denote the step-wise reward, discount factor, and horizon. To bridge this long-term objective with immediate decisions, the action-value function (Q-function) is:

$$Q(s_t, a_t; \mathcal{I}^*) = \mathbb{E} \left[\sum_{k=t}^T \gamma^{k-t} R_k(s_k, a_k; \mathcal{I}^*) \mid s_t, a_t, \mathcal{I}^* \right]. \quad (3)$$

This component represents the operator's "perfect knowledge" of the task requirements.

2) Bounded Decision-Making: While the objective Q is optimal, the human's ability to maximize it is constrained. To capture bounded rationality, we assume the human's actual control input $a_t^H \in \mathcal{A}$ follows a *stochastic policy* derived from Q-values:

$$a_t^H \sim \pi^H(a_t^H \mid s_t; \mathcal{I}^*), \quad (4)$$

which models bounded rationality as noisy decision-making governed by two essential properties:

- (i) **Intention-Consistency:** Higher probability is assigned to actions yielding higher Q-values, reflecting the goal-directed nature:

$$\begin{aligned} Q(s_t, a_1; \mathcal{I}^*) &> Q(s_t, a_2; \mathcal{I}^*) \\ \implies \pi^H(a_1 \mid s_t; \mathcal{I}^*) &\geq \pi^H(a_2 \mid s_t; \mathcal{I}^*). \end{aligned}$$

- (ii) **Full Support:** $\pi^H(a_t^H \mid \cdot) > 0$ for all feasible actions, accounting for potential execution noise and exploration.

This formulation aligns with the theory of reward-rational choice [41], which provides a mathematical abstraction of bounded-rational, goal-directed behavior. By representing human commands as stochastic but intention-dependent actions, Eq. (4) captures a broad class of suboptimal behaviors in which actions remain statistically related to the underlying intention. Standard models, including the widely used Boltzmann rationality [14], can be viewed as particular instances of this abstraction. Because Eq. (4) defines a conditional distribution of human commands given the state and intention, it can also serve as a predictive human-input model.

Remark 1: The stochastic-policy model should be interpreted as an abstract probabilistic description of the action–intention relationship, rather than a complete psychological or physiological model of human suboptimality. Therefore, the model is most appropriate when human commands remain statistically informative about the intended goal. However, factors such as fatigue, delayed response, or persistent bias may change this relationship by making the observed commands noisier, delayed, or systematically shifted. If the commands become weakly informative or misleading, the model may become mismatched and its predictive reliability may decrease.

- **Autonomous Agent (Partial Information, Perfect Rationality):** The autonomous agent is modeled as a

perfectly rational entity but operates under partial information regarding $\mathcal{I}^*$. Its process consists of: **1) Intention Inference:** The inference module maintains a belief distribution $b_t(\mathcal{I})$ at time t :

$$b_t(\mathcal{I}) = P(\mathcal{I} \mid a_{0:t}^H, s_{0:t}). \quad (5)$$

From a dynamical system perspective, this inference process can be abstracted as a belief transition mapping $b_{t+1} = \Psi(b_t, s_{t+1}, a_{t+1}^H)$. Here, the intention space Θ is assumed to be finite (or pre-defined) and known to the agent. The maximum a posteriori (MAP) estimate $\hat{\mathcal{I}}_t = \arg \max_{\mathcal{I} \in \Theta} b_t(\mathcal{I})$ is used as the current estimate.

2) Rational Controller: Based on $\hat{\mathcal{I}}_t$, the agent generates assistive actions. To focus the formulation on authority allocation under intention uncertainty induced by bounded-rational human inputs, we assume that the autonomous agent can generate an optimal assistive action for a specified intention. This idealized assumption abstracts away low-level controller design and is approximated in the experiments by the task-specific autonomy modules described in Section V. Driven by this assumption of perfect rationality, the agent selects:

$$a_t^A = \arg \max_{a_t \in \mathcal{A}} Q(s_t, a_t; \hat{\mathcal{I}}_t), \quad (6)$$

where a_t^A shares the same action space $\mathcal{A}$ as the human operator.

Remark 2: It is worth noting that the agent plans based on the most likely intention (MAP) rather than the expected Q-value over the belief distribution. This design is deliberate: (1) In spatial tasks, averaging policies over conflicting intentions can lead to valid but unsafe trajectories (e.g., steering into obstacles); (2) It strictly separates the role of execution (Agent) from uncertainty management (Arbitrator). The Agent provides a coherent optimal candidate action, while the Arbitrator regulates its influence based on uncertainty levels.

B. Problem and Objective

The mathematical modeling above shows that authority allocation is not merely a static fusion problem, because the current authority decision affects both immediate task execution and future intention inference. This coupling leads to two key challenges that motivate an information-aware optimization framework:

1) Propagation of Intention Uncertainty: The bounded-rational human model in Eq. (4) implies that human actions are stochastic samples $a_t^H \sim \pi^H(\cdot \mid s_t; \mathcal{I}^*)$, so a single input a_t^H cannot deterministically reveal the true intention $\mathcal{I}^*$. This stochasticity propagates through the intention-inference update in Eq. (5), resulting in a belief distribution b_t that may retain substantial uncertainty after limited observations. Because the autonomous action in Eq. (6) is generated according to the inferred intention $\hat{\mathcal{I}}_t$, uncertainty in b_t increases the risk of inappropriate assistance when the inferred intention is incorrect or insufficiently reliable. Therefore, a central challenge is to reduce intention uncertainty efficiently while maintaining task-oriented control.

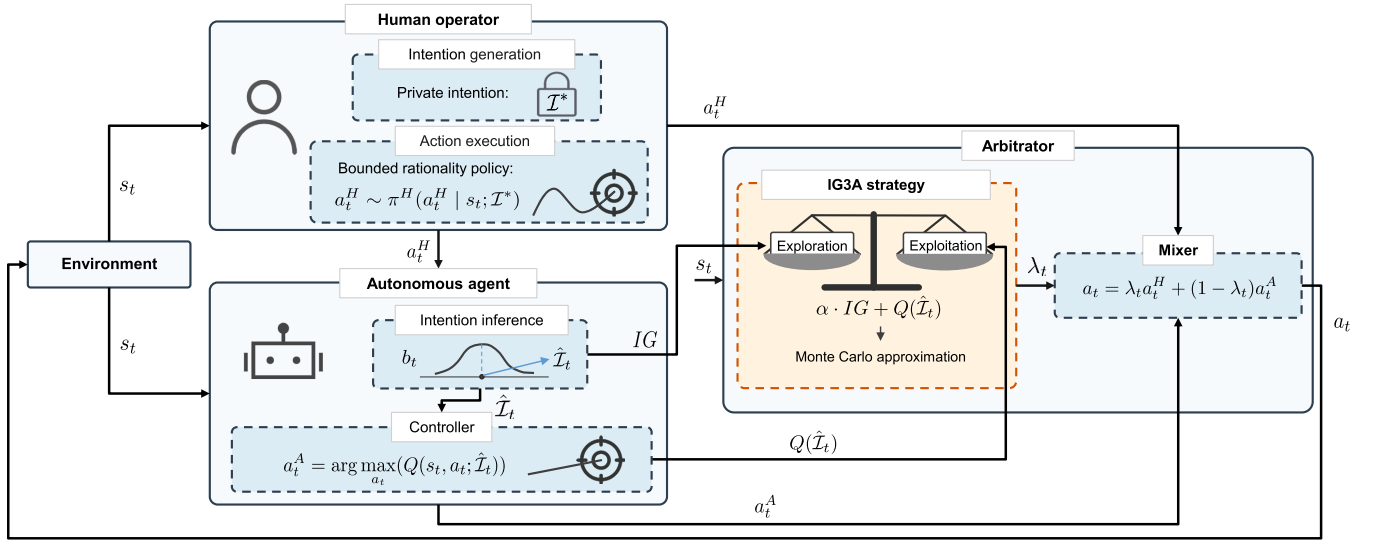


Fig. 2. Overview of the proposed IG3A framework for shared control under information-decision asymmetry. The core Arbitrator employs the IG3A strategy to dynamically calculate the authority λ_t , explicitly balancing the trade-off between task exploitation (Q) and intention exploration (IG) to actively reduce intention uncertainty.

2) Information–Decision Coupling in Authority Allocation:

A purely reactive authority rule, such as a static mapping $\lambda_t = f(b_t)$ based only on the current belief or confidence, does not explicitly account for how the current authority decision shapes future observations. The recursive belief dynamics reveal this coupling: the current authority λ_t affects the executed action and the next state s_{t+1} , which in turn influences the future human response a_{t+1}^H and the updated belief b_{t+1} . Thus, authority should be viewed not only as a passive fusion weight for the current inputs, but also as a decision variable with an active influence on future intention information. An ideal long-horizon formulation would therefore select a sequence of authority coefficients $\lambda_{1:T}$ by considering both task return and the belief evolution induced by future human–machine interactions:

$$\begin{aligned}
 & \max_{\lambda_{1:T}} \mathbb{E}_{\mathcal{I} \sim b_t} \left[\sum_{k=t}^T \gamma^{k-t} R_k(s_k, a_k; \mathcal{I}) \right] \\
 \text{s.t.} \quad & s_{k+1} \sim p(s_{k+1} | s_k, a_k), & (\text{Environment}) \\
 & a_k = \lambda_k a_k^H + (1 - \lambda_k) a_k^A, & (\text{Arbitration}) \\
 & a_k^H \sim \pi^H(\cdot | s_k; \mathcal{I}), & (\text{Human}) \\
 & a_k^A = \arg \max_{a_t} Q(\cdot; \hat{\mathcal{I}}_k), & (\text{Agent}) \\
 & b_{k+1} = \Psi(b_k, s_{k+1}, a_{k+1}^H), & (\text{Inference}) \\
 & \lambda_k \in [0, 1].
 \end{aligned} \tag{7}$$

Directly solving Eq. (7) in real time is computationally intractable, which motivates the tractable surrogate formulation introduced in the next section.

IV. METHODOLOGY OF IG3A FRAMEWORK

As illustrated in Fig. 2, this section presents the core methodology of the IG3A framework to solve the optimization problem formulated in Section III. Building upon the ideal modeling of information-decision asymmetry, we propose an active authority allocation strategy that strategically regulates control to facilitate intention expression.

The core innovation lies in the solution design: recognizing that the ideal policy (Eq. 7) is computationally intractable,

we construct a tractable surrogate objective. This dual-objective formulation—central to the IG3A strategy—enables the autonomous agent to dynamically balance task exploitation with intention exploration, actively reducing uncertainty while ensuring performance.

In the following, we first define this surrogate objective and then introduce an efficient online solver based on discretized authority evaluation and Monte Carlo approximation for real-time deployment.

A. Dual-Objective With Information Gain

From a belief-space optimal control perspective, the authority coefficient λ_t affects not only the immediate task performance but also the evolution of the belief over human intentions. This dual influence is consistent with both the classical dual control principle [42] and the dual effect formulation [43], where control actions simultaneously contribute to system regulation and information acquisition. Accordingly, the decision-making objective naturally involves two aspects:

- **Exploitation:** Ensuring high task performance at the current time step.
- **Exploration:** Actively reducing uncertainty in human intention to accelerate belief convergence.

Motivated by these considerations, the following surrogate value function is proposed to enable real-time shared control:

$$Q_{\text{sur}}(b_t, s_t, \lambda_t) = \underbrace{Q(s_t, a_t(\lambda_t); \hat{\mathcal{I}}_t)}_{\text{exploitation term}} + \underbrace{\alpha \cdot IG(\lambda_t)}_{\text{exploration term}}, \tag{8}$$

where $IG(\lambda_t)$ quantifies the expected information gain from executing λ_t as defined in Eq. (9), and $\alpha > 0$ is the exploration weight for balancing exploration–exploitation trade-offs. This surrogate objective is introduced to enable tractable decision-making for the original belief-space optimization problem, rather than as an exact or globally optimal reformulation, while remaining consistent with the dual role of authority allocation under uncertainty.

Specifically, the information gain $IG(\lambda_t)$ is defined as:

$$IG(\lambda_t) = H(b_t) - \mathbb{E}[H(b_{t+1} | \lambda_t)], \quad (9)$$

where $H(b_t)$ denotes the Shannon entropy [44] of belief b_t , defined as

$$H(b_t) = - \sum_{\mathcal{I} \in \Theta} b_t(\mathcal{I}) \log b_t(\mathcal{I}). \quad (10)$$

The expected posterior entropy $\mathbb{E}[H(b_{t+1} | \lambda_t)]$, representing the expected uncertainty after authority allocation λ_t , is defined as:

$$\mathbb{E}[H(b_{t+1} | \lambda_t)] \triangleq \mathbb{E}_{a_{t+1}^H, s_{t+1} \sim p(\cdot | b_t, s_t, \lambda_t)} [H(b_{t+1})]. \quad (11)$$

Intuitively, $IG(\lambda_t)$ quantifies the expected reduction in uncertainty about the human's intention resulting from the selected authority allocation. From a broader perspective, this formulation can be viewed as a specific instantiation of a more general class of exploration measures. In principle, the exploration component can be defined using any scalar function that quantifies the uncertainty of the intention belief and its expected reduction under a given authority allocation. In this work, entropy-based information gain is adopted due to its compatibility with the belief representation, clear information-theoretic interpretation, and computational tractability, making it suitable for real-time shared control. Accordingly, the overall objective can be understood as a myopic, structure-preserving surrogate.

The expected posterior entropy is derived by marginalizing over future observations. Starting from the law of total expectation:

$$\begin{aligned} \mathbb{E}[H(b_{t+1} | \lambda_t)] &= \int_{\mathcal{S}} \int_{\mathcal{A}^H} \underbrace{H(b_{t+1})}_{\text{entropy of posterior}} \cdot \underbrace{p(a_{t+1}^H, s_{t+1} | b_t, s_t, \lambda_t)}_{\text{predictive distribution}} da_{t+1}^H ds_{t+1} \\ &= \int_{\mathcal{S}} \int_{\mathcal{A}^H} H(b_{t+1}) \cdot p(a_{t+1}^H | s_{t+1}, b_t) \cdot p(s_{t+1} | \lambda_t) da_{t+1}^H ds_{t+1} \\ &= \int_{\mathcal{S}} \int_{\mathcal{A}^H} H(b_{t+1}) \cdot \left[\sum_{\mathcal{I}'} p(a_{t+1}^H | s_{t+1}, \mathcal{I}') b_t(\mathcal{I}') \right] \cdot p(s_{t+1} | s_t, a_t(\lambda_t)) da_{t+1}^H ds_{t+1}, \end{aligned} \quad (12)$$

where:

- $p(a_{t+1}^H | s_{t+1}, b_t) = \sum_{\mathcal{I}'} p(a_{t+1}^H | s_{t+1}, \mathcal{I}') b_t(\mathcal{I}')$ is the predictive human action distribution [Eq. (4)].
- $p(s_{t+1} | s_t, a_t(\lambda_t))$ is the system dynamics.

This decomposition enables tractable estimation of $\mathbb{E}[H(b_{t+1} | \lambda_t)]$ for any $\lambda_t \in [0, 1]$, allowing the information-exploration term to be computed efficiently in real time.

It is important to note that the efficacy of the information gain term relies fundamentally on the human behavioral assumptions established in Section III. Specifically, the Intention-Consistency property ensures that human actions

statistically reflect their underlying goals, creating the “signal” that the inference module detects. Meanwhile, the Full Support property necessitates the probabilistic treatment of inputs, creating the “uncertainty” that IG3A seeks to reduce. Although this surrogate does not guarantee monotonic entropy reduction along every realized trajectory, the information-gain term favors authority choices that are expected, under the predictive model, to yield lower posterior uncertainty.

With this surrogate objective, real-time optimization becomes feasible, while still maintaining the critical balance between exploitation and information exploration. The surrogate objective has the following interpretations:

- When b_t is concentrated around the current true intention $\mathcal{I}^*$, the surrogate objective is dominated by the exploitation term, and its authority preference tends to align with the long-horizon authority-value evaluation in Eq. (7).
- When b_t is uncertain, the information-exploration term captures the future-exploration value implicit in Eq. (7) by favoring authority choices with lower expected posterior entropy.
- The surrogate avoids long-horizon belief-space optimization by combining single-step task evaluation with low-dimensional information estimation, enabling real-time implementation while preserving the exploration-exploitation trade-off.

B. Practical Solution to the Optimization Problem

Although the expected posterior entropy $\mathbb{E}[H(b_{t+1} | \lambda_t)]$ can be explicitly written as a nested integral over future states and human actions as in Eq. (12), its evaluation requires high-dimensional marginalization, which is computationally prohibitive in practice.

To address this, a Monte Carlo sampling approach is employed to approximate it. Specifically, the continuous authority parameter $\lambda_t \in [0, 1]$ is discretized into a finite set of values at fixed intervals, which significantly simplifies the computation. The information gain of each candidate λ_t is computed using Alg. 1. In this manner, an otherwise intractable optimization problem is transformed into a computationally feasible one through forward simulation and sampling, thereby enabling practical deployment in real-world robotic systems.

Based on Algorithm 1, the computational cost is dominated by evaluating all candidate authority levels and estimating the expected posterior entropy via Monte Carlo sampling. Let $M = |\Lambda|$ be the number of candidate authority levels, N the number of samples, and $K = |\Theta|$ the number of intention hypotheses. For each $\lambda_t \in \Lambda$, the algorithm performs N one-step state predictions and N posterior updates with entropy evaluations, yielding an overall complexity

$$O(MN(C_{\text{dyn}} + K)). \quad (13)$$

Here, C_{dyn} denotes the cost of one-step predictive state propagation, which may be implemented by an analytical dynamics model, a learned predictive model, or any other sampling-capable forward model. The term K arises from belief updating and entropy computation over the intention space, and the

Algorithm 1 Monte Carlo Approximation for Surrogate-Based Authority Selection

Require: Current belief $b_t(\mathcal{I})$, current state s_t , candidate authority levels $\Lambda = \{\lambda^{(1)}, \dots, \lambda^{(M)}\} \subset [0, 1]$, number of Monte Carlo samples N , balancing weight of exploitation and exploration α ;

Ensure: Selected authority λ_t^* .

```

1: for each candidate  $\lambda_t \in \Lambda$  do
2:   Step 1: Predict future states
3:   Sample  $\{s_{t+1}^{(i)}\}_{i=1}^N \sim p(s_{t+1} | s_t, a_t(\lambda_t))$ 
4:   Initialize empty list  $\mathcal{H}_{\text{samples}} \leftarrow []$ 
5:   Step 2: Simulate human responses and belief updates
6:   for  $i = 1$  to  $N$  do
7:     Sample  $\mathcal{I}^{(i)} \sim b_t(\mathcal{I})$   $\triangleright$  Sample intention from current belief
8:     Sample  $a_{t+1}^{H(i)} \sim p(a_{t+1}^H | s_{t+1}^{(i)}, \mathcal{I}^{(i)})$ 
9:     Compute updated posterior:
10:     $b_{t+1}^{(i)}(\mathcal{I}) \leftarrow P(\mathcal{I} | a_{0:t}^H, s_{0:t}, a_{t+1}^{H(i)}, s_{t+1}^{(i)})$ 
11:    Compute entropy of the updated posterior:
12:     $H(b_{t+1}^{(i)}) \leftarrow -\sum_{\mathcal{I}'} b_{t+1}^{(i)}(\mathcal{I}') \log b_{t+1}^{(i)}(\mathcal{I}')$ 
13:    Append  $H(b_{t+1}^{(i)})$  to  $\mathcal{H}_{\text{samples}}$ 
14:   end for
15:   Step 3: Estimate expected posterior entropy
16:   Compute the mean of the sampled entropies:
17:    $\mathbb{E}[H(b_{t+1} | \lambda_t)] \approx \frac{1}{N} \sum_{i=1}^N H(b_{t+1}^{(i)})$ 
18:   Step 4: Evaluate objective
19:   Compute  $Q_{\text{sur}}(b_t, s_t, \lambda_t) = Q(s_t, a_t(\lambda_t); \hat{\mathcal{I}}_t) + \alpha \cdot [H(b_t) - \mathbb{E}[H(b_{t+1} | \lambda_t)]]$ 
20: end for
21: Step 5: Select authority
22: Select  $\lambda_t^* \leftarrow \arg \max_{\lambda_t \in \Lambda} Q_{\text{sur}}(b_t, s_t, \lambda_t)$ 
23: return  $\lambda_t^*$ 

```

overall complexity is linear in both the number of candidate intentions and the Monte Carlo sample size. In typical shared-control settings, the intention space is task-defined and of moderate size (≤ 20), so this term remains manageable in practice. Therefore, aside from the cost of the chosen forward model, the proposed authority evaluation scales linearly with the number of candidate authority levels, Monte Carlo samples, and intention hypotheses.

V. EXPERIMENT

To assess the effectiveness of the proposed method, experiments were conducted in simulated-human LunarLander tasks, human-in-the-loop (HITL) LunarLander tasks, and HITL physical robotic-arm experiments. The simulated-human experiments serve as the main controlled quantitative testbed, enabling matched-seed repeated evaluations under intention switching, disturbances, degraded human commands, and parameter variations. The HITL experiments in LunarLander and physical robotic-arm provide complementary validation with real human inputs and hardware execution.

A. Common Experimental Setup

1) *Compared Control Methods:* Four representative control methods are evaluated in the main experiments:

- **Confidence-based Shared Control (CB).** CB represents a classical belief-reactive arbitration strategy, where the human authority is determined solely by the confidence of the current intention estimate [28]. It is included to examine whether active intention exploration provides benefits beyond conventional confidence-based authority allocation. The human authority is computed as

$$\lambda(t) = \begin{cases} \delta_1, & C(t) \leq \sigma_1, \\ \frac{\delta_1 - \delta_2}{\sigma_1 - \sigma_2} C(t) + \frac{\delta_1 \sigma_2 - \delta_2 \sigma_1}{\sigma_2 - \sigma_1}, & \sigma_1 < C(t) \leq \sigma_2, \\ \delta_2, & C(t) > \sigma_2, \end{cases} \quad (14)$$

where $C(t)$ denotes the confidence margin of the inferred intention:

$$C(t) = \max_{\mathcal{I} \in \Theta} b_t(\mathcal{I}) - \max_{\mathcal{I} \in \Theta \setminus \{\hat{\mathcal{I}}_t\}} b_t(\mathcal{I}). \quad (15)$$

- **Hindsight Optimization (HO).** HO is included as a representative POMDP-style shared-autonomy baseline under goal uncertainty [16]. It evaluates assistance by considering possible goals under the current belief and approximates the intractable POMDP solution through hindsight optimization, thereby providing a strong belief-aware comparison for IG3A.
- **Interventional Diffusion Assistance (IDA).** IDA is included as a recent learning-based shared-autonomy baseline [18]. It uses a diffusion-based copilot with a goal-agnostic intervention rule to determine when assistance should override or modify the human input. Unlike CB, HO, and IG3A, IDA does not explicitly use the shared discrete intention-belief module, making it a complementary data-driven comparison.
- **IG3A (Proposed).** IG3A is the proposed active authority-allocation method, which selects the authority coefficient by optimizing the surrogate objective in Eq. (8).

Additional ablation and reference baselines, including human-only control, fixed authority allocation, and the pure information-gain variant, are provided in the Supplementary Material.

2) *Human Decision Model, Intention Inference, and Q-Function Implementation:* To isolate differences in the shared-control strategy from differences in intention recognition, CB, HO, and IG3A use the same probabilistic intention-inference module [14]. IDA is treated as a learning-based baseline and does not use this explicit intention-inference module.

The shared module maintains a belief distribution $b_t(\mathcal{I})$ over the discrete intention set Θ , which is recursively updated as

$$b_t(\mathcal{I}) \propto P(a_t^H | s_t, \mathcal{I}) \sum_{\mathcal{I}' \in \Theta} P(\mathcal{I} | \mathcal{I}') b_{t-1}(\mathcal{I}'). \quad (16)$$

The action likelihood is implemented using a Boltzmann-rational model,

$$P(a_t^H | s_t, \mathcal{I}) \propto \exp(\hat{\beta} Q(s_t, a_t^H; \mathcal{I})), \quad (17)$$

where $\hat{\beta}$ denotes the rationality parameter assumed by the inference model and can be estimated from human demonstration data by maximum likelihood [14]. The same Boltzmann

form is also used as the predictive human-input model in the Monte Carlo estimation of the expected posterior entropy, where future human responses are sampled under candidate intentions.

In the experiments, the action-value function $Q(s_t, a; \mathcal{I})$ used in both the likelihood model and the exploitation term is implemented by the quadratic surrogate

$$Q(s_t, a; \mathcal{I}) = -\|a - a_t^A(\mathcal{I})\|^2,$$

where $a_t^A(\mathcal{I})$ denotes the autonomy action associated with intention $\mathcal{I}$. Thus, Eq. (17) evaluates this surrogate at $a = a_t^H$, while Eq. (8) evaluates it at the blended action $a = a_t(\lambda_t)$ under the current intention estimate.

The intention transition model is

$$P(\mathcal{I}_i | \mathcal{I}_j) = \begin{cases} 1 - \Delta, & i = j, \\ \frac{\Delta}{|\Theta| - 1}, & i \neq j, \end{cases} \quad (18)$$

where Δ denotes the assumed per-step switching probability. Environment-specific values of $\hat{\beta}$, Δ , and other implementation parameters are provided in the corresponding experiment sections.

3) Evaluation and Statistical Reporting: In simulated-human experiments, matched random seeds were used across all compared methods. Each seed defined a matched task instance, including the initial condition, disturbance realization, and target configuration. Because residual stochasticity remained due to simulated human inputs, Monte Carlo sampling, and stochastic policy execution, each seed was evaluated with multiple repeated runs. For each method and condition, repeated runs were first averaged within each seed, yielding one seed-level estimate per method. Curves and tables report the mean across seed-level estimates, with error bars or intervals corresponding to $1.96 SE$, where SE is the standard error across seeds. For paired comparisons, differences were computed using the seed-level estimates, which preserves the matched-task structure across methods and avoids treating repeated runs under the same seed as independent samples. Statistical significance was assessed using the Wilcoxon signed-rank test on the seed-level paired differences, with significance levels denoted as $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$.

For human-participant experiments, repeated trials were first averaged within each participant and condition, yielding one participant-level estimate per method. Figures report the distribution of participant-level estimates, where boxes summarize the interquartile range and median, dots denote individual participants, and horizontal bars with significance markers indicate Wilcoxon signed-rank tests on participant-level paired differences between the indicated method pairs. This reporting protocol avoids treating repeated trials from the same participant as independent samples.

B. Simulated-Human LunarLander Experiments

1) Environment and Implementation: The simulations are based on a customized Gymnasium environment [45], derived from *LunarLanderContinuous-v2*, in which the lander must safely touch down within a designated target zone (Fig. 3).

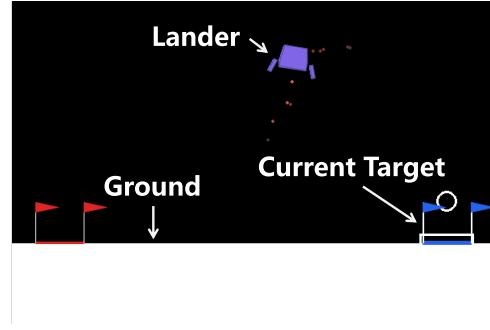


Fig. 3. Customized LunarLander environment with multiple candidate landing zones.

Randomized initial impulses, wind forces, and torque perturbations are applied in the environment. Multiple candidate landing zones define the discrete intention set shared by the human operator and the autonomous agent; unless otherwise stated, the default 2-Spread setting contains two zones centered at $x = -0.8$ and $x = 0.8$. In switching episodes, the simulated human intention changes between candidate zones, with the pre-switch and post-switch targets sampled according to the matched random seed. Additional goal configurations are examined later to evaluate robustness to task variations.

The lander is actuated by a main vertical engine and two lateral engines, and both the human operator and the autonomous agent generate continuous commands from the 8-dimensional state, including position, velocity, angle, angular velocity, and leg-contact indicators. For each candidate intention, the autonomous action is generated by an offline-trained SAC policy, serving as an approximate implementation of the idealized rational controller in Eq. (6); this intention-conditioned policy set is fixed across all compared methods. Simulated human commands are sampled from the Boltzmann-rational model in Eq. (17) with rationality parameter β_H ; except in the model-mismatch evaluation, the inference parameter is set to $\hat{\beta} = \beta_H$.

A trial is considered successful if the lander touches down smoothly within a tolerance of ± 0.2 around the current target zone. The reward follows the original LunarLander shaping structure, with terms for target distance, velocity, attitude angle, fuel consumption, leg contact, and successful landing. For intention-switching episodes, the shaping baseline is reset when the target changes, preventing artificial reward jumps caused solely by replacing the old target with the new one. Terminal success or failure is always evaluated with respect to the current target.

The intention-transition probability in Eq. (18) is set to $\Delta = 1/300$, and the CB parameters in Eq. (14) are set to $\delta_1 = 0.7$, $\delta_2 = 0.1$, $\sigma_1 = 0.2$, and $\sigma_2 = 0.8$. For IG3A, λ is discretized over $[0.1, 0.7]$ with step size 0.05, resulting in 13 candidates; the number of particles is set to 30, and the information-gain weight is set to $\alpha = 30$. These values are chosen based on preliminary tuning and are further examined in the parameter sensitivity analysis. Unless otherwise specified, each simulation condition is evaluated with 50 matched seeds and 10 repeated runs per seed. The one-step state prediction in line 3 of Alg. 1 is obtained by querying

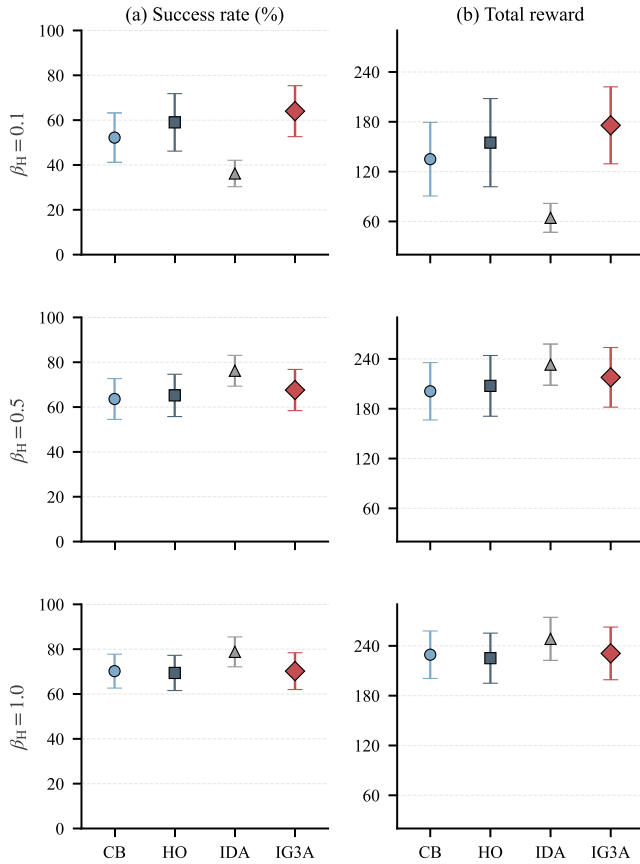


Fig. 4. Performance under intention switching in the 2-Spread setting. (a) Success rate. (b) Total reward. Rows correspond to $\beta_H = 0.1, 0.5$, and 1.0 from top to bottom.

the customized simulator as a sampling-based forward model, which requires no explicit analytical dynamics equation and follows the general forward-model interface assumed in Alg. 1.

2) *Main Performance Under Intention Switching*: The main simulation study evaluates the four compared methods under intention switching, where the simulated human intention switches once at step 100. Unlike fixed-intention episodes, previously accumulated intention evidence may become less relevant after the switch, requiring the shared-control strategy to recover reliable assistance from bounded-rational inputs. This setting therefore stresses the ability of each method to handle noisy, weak, or temporarily inconsistent intention evidence.

Fig. 4 shows that IG3A achieves the highest success rate and total reward under the low-rationality condition $\beta_H = 0.1$. The seed-level paired comparisons in Fig. 5 further show that these improvements are significant against all three baselines in both metrics. Together, these results indicate that, under weak intention evidence, explicitly optimizing the trade-off between task exploitation and active intention exploration can be advantageous compared with relying only on the current belief estimate or action-level intervention.

As β_H increases, human actions become more goal-consistent, the paired differences become smaller, and the advantage of active authority allocation is no longer uniform.

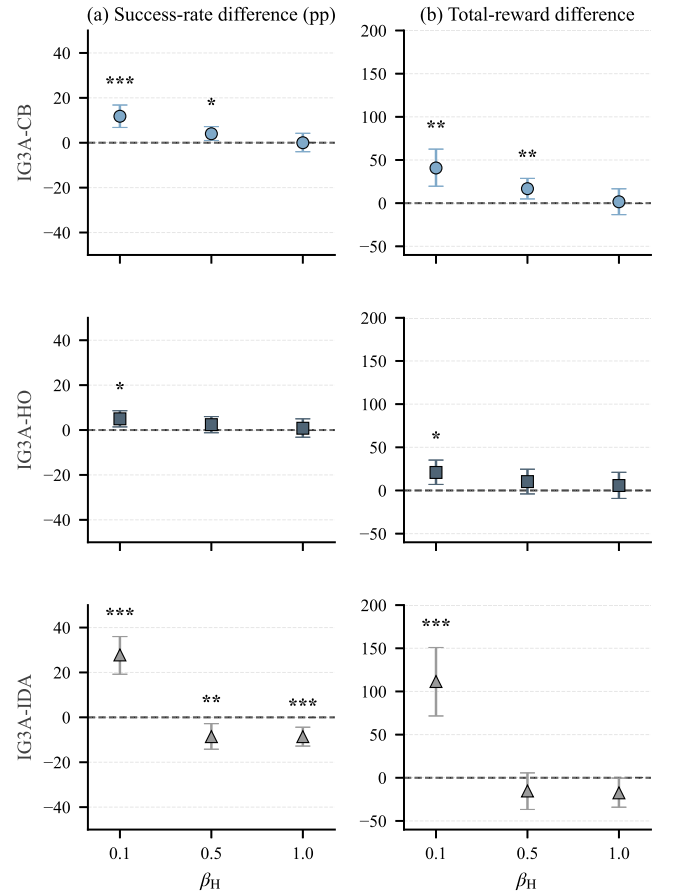


Fig. 5. Seed-level paired differences between IG3A and each baseline under intention switching in the 2-Spread setting. (a) Success-rate difference. (b) Total-reward difference. Rows correspond to comparisons with CB, HO, and IDA from top to bottom. Positive values indicate better performance of IG3A. Significance markers follow the convention defined in Section V-A3. Success-rate differences are reported in percentage points (pp).

For $\beta_H = 0.5$ and $\beta_H = 1.0$, IDA obtains slightly higher success rates, which is consistent with its action-level intervention design: when human commands are already informative, direct intervention can be sufficient for successful landing. However, in these settings, IDA's higher success rate does not translate into a clear reward advantage, and IG3A remains comparable in total reward. Overall, IG3A shows its clearest task-level advantage in the difficult low-rationality switching condition, while remaining comparable in total reward when the human commands become more goal-consistent.

Table I reports the belief convergence and post-switch recovery steps in the 2-Spread intention-switching setting, which are used to diagnose how quickly each method provides evidence for the currently assigned intention. Since IDA does not maintain an explicit intention-inference module, its convergence and recovery steps are computed using the same diagnostic belief estimator as the other methods, but this estimate does not affect IDA's executed actions. IG3A achieves the fastest or near-fastest initial convergence across the tested rationality levels, suggesting that the information-gain term is effective when the belief is highly uncertain. For post-switch recovery, the required steps decrease for all methods as β_H increases; IG3A is not the fastest under severely degraded

TABLE I

SEED-LEVEL CONVERGENCE AND RECOVERY STEPS UNDER INTENTION SWITCHING IN THE 2-SPREAD SETTING. CONVERGENCE IS MEASURED AS THE FIRST STEP AT WHICH THE BELIEF OF THE INITIALLY ASSIGNED TARGET REACHES 0.9, WHEREAS RECOVERY IS MEASURED AS THE FIRST POST-SWITCH STEP AT WHICH THE BELIEF OF THE NEW TARGET REACHES 0.9. VALUES ARE MEAN $\pm 1.96 SE$, WITH LOWER VALUES INDICATING FASTER BELIEF ADAPTATION

(a) Convergence time (steps)				
β_H	CB	HO	IDA	IG3A
0.1	21.40 $\pm$ 2.22	16.19 $\pm$ 1.14	22.91 $\pm$ 1.93	20.11 $\pm$ 1.57
0.5	3.60 $\pm$ 0.33	3.42 $\pm$ 0.39	3.48 $\pm$ 0.26	3.03 $\pm$ 0.27
1.0	1.90 $\pm$ 0.12	1.90 $\pm$ 0.13	1.98 $\pm$ 0.15	1.84 $\pm$ 0.18

(b) Recovery time (steps)				
β_H	CB	HO	IDA	IG3A
0.1	115.18 $\pm$ 13.91	118.62 $\pm$ 11.72	54.38 $\pm$ 3.97	122.66 $\pm$ 13.11
0.5	15.72 $\pm$ 0.83	17.53 $\pm$ 0.89	7.80 $\pm$ 0.50	16.67 $\pm$ 1.28
1.0	7.71 $\pm$ 0.49	8.71 $\pm$ 0.47	4.97 $\pm$ 0.42	7.65 $\pm$ 0.48

TABLE II

SENSITIVITY TO GOAL CONFIGURATION AND INTENTION DYNAMICS UNDER LOW-RATIONALITY SIMULATED HUMANS ($\beta_H = 0.1$). VALUES ARE SUCCESS RATES (%) REPORTED AS MEAN $\pm 1.96 SE$

Goals	Dynamics	CB	HO	IDA	IG3A
2-Spread	No switch	95.6 $\pm$ 2.3	98.8 $\pm$ 1.3	64.0 $\pm$ 6.4	98.0 $\pm$ 1.7
3-Close	Switch	81.2 $\pm$ 5.9	95.8 $\pm$ 2.9	61.8 $\pm$ 4.2	91.4 $\pm$ 3.3
3-Spread	Switch	83.6 $\pm$ 6.5	93.8 $\pm$ 5.6	47.2 $\pm$ 4.9	92.0 $\pm$ 4.6

inputs, but becomes comparable to or slightly better than CB and HO under moderate and high rationality. This result indicates that IG3A does not simply pursue the fastest belief recovery, but balances intention exploration with immediate task execution, so recovery time should be interpreted together with success rate and total reward.

3) *Robustness Under Task and Interaction Variations:* This subsection evaluates the robustness of IG3A under three non-ideal variations beyond the main intention-switching experiment: task-level changes, imperfect human commands, and noisy state observations.

a) *Goal configuration and intention dynamics:* To examine whether the observations above depend on a specific goal configuration or intention-switching condition, additional simulations were conducted under different goal configurations and intention dynamics. Table II summarizes the results for $\beta_H = 0.1$, since this condition produces the clearest separation among methods. The 3-Close and 3-Spread settings use target-center x -coordinates $(-0.4, 0, 0.4)$ and $(-0.8, 0, 0.8)$, respectively. Results for larger β_H are provided in the Supplementary Material, where the differences among methods are generally smaller because the human inputs become more informative.

Across these additional cases, IG3A remains among the top-performing methods and substantially outperforms IDA under low-rationality inputs. In the no-switch two-goal case, HO and

TABLE III

PERFORMANCE UNDER DEGRADED SIMULATED HUMANS ON 2-SPREAD WITH INTENTION SWITCHING ($\beta_H = 0.5$). VALUES ARE MEAN $\pm 1.96 SE$. FOR EACH VARIANT AND METRIC, THE BEST RESULT IS SHOWN IN **BOLD**, AND THE SECOND-BEST IS UNDERLINED. "DISCONN." DENOTES DISCONNECTING

(a) Success rate (%)				
Variant	CB	HO	IDA	IG3A
Laggy	69.2 $\pm$ 7.4	73.6$\pm$8.3	19.2 $\pm$ 4.7	<u>70.0$\pm$6.3</u>
Biased	56.4 $\pm$ 12.6	56.2 $\pm$ 13.2	68.2$\pm$9.3	<u>57.0$\pm$13.0</u>
Disconn.	56.8 $\pm$ 11.2	<u>57.6$\pm$11.2</u>	56.0 $\pm$ 11.3	59.0$\pm$11.3
Random	53.6 $\pm$ 11.3	<u>56.8$\pm$11.6</u>	<u>56.8$\pm$9.3</u>	57.2$\pm$10.9

(b) Reward				
Variant	CB	HO	IDA	IG3A
Laggy	201.4 $\pm$ 29.6	218.9$\pm$32.7	-8.5 $\pm$ 20.2	<u>203.9$\pm$24.0</u>
Biased	173.2 $\pm$ 49.6	168.9 $\pm$ 54.1	195.6$\pm$33.5	<u>174.2$\pm$51.8</u>
Disconn.	173.7 $\pm$ 43.1	<u>175.1$\pm$44.1</u>	141.3 $\pm$ 39.8	180.3$\pm$43.7
Random	158.7 $\pm$ 43.7	<u>171.6$\pm$45.6</u>	156.6 $\pm$ 30.2	172.3$\pm$42.7

IG3A both achieve near-ceiling success rates. In the three-goal switching cases, HO achieves slightly higher success rates, while IG3A remains close to HO, suggesting robustness across different goal geometries and intention dynamics.

b) *Imperfect human commands:* To evaluate robustness beyond the nominal Boltzmann-rational input model, we test the four compared methods under four degraded operator variants in the 2-Spread intention-switching setting with $\beta_H = 0.5$. Starting from the nominal Boltzmann-rational simulated human, Laggy repeats the previous action with probability 0.85; Biased adds a constant lateral offset of -0.3 to the horizontal input; Disconnecting outputs a zero action with probability 0.3; and Random replaces the intended action with a uniformly sampled action with probability 0.3.

Table III shows that IG3A remains competitive across all four imperfect operator types. Although the best-performing method varies with the degradation pattern, IG3A consistently remains among the top-performing methods and avoids the large performance drops observed for some baselines, suggesting stable performance under moderate human-input model mismatch.

c) *Noisy state observations:* To examine robustness to observation or state-estimation errors, the compared methods are tested under noisy observations in the same 2-Spread intention-switching setting as the main simulation study. Gaussian observation noise with normalized standard deviation σ_{obs} is injected into the continuous state dimensions, while the discrete leg-contact indicators are kept unchanged.

Table IV shows that IG3A remains competitive under noisy observations, outperforming CB and IDA across all tested noise levels and remaining close to HO when the noise level increases. These results indicate that IG3A can tolerate moderate observation-side imperfections, although HO is slightly more robust under larger observation noise in this setting.

4) *Parameter and Computational Sensitivity:* We examine the sensitivity of IG3A to key design and computational parameters, including the exploration weight α , rationality-model mismatch between β_H and $\hat{\beta}$, Monte Carlo sampling

TABLE IV

SENSITIVITY TO NORMALIZED OBSERVATION NOISE UNDER LOW-RATIONALITY SIMULATED HUMANS ($\beta_H = 0.1$). VALUES ARE SUCCESS RATES (%) REPORTED AS MEAN $\pm 1.96SE$. THE NOISE LEVEL σ_{obs} DENOTES THE NORMALIZED STANDARD DEVIATION OF GAUSSIAN NOISE APPLIED TO CONTINUOUS OBSERVATION DIMENSIONS, WITH DISCRETE LEG-CONTACT INDICATORS EXCLUDED

σ_{obs}	CB	HO	IDA	IG3A
0.05	44.8 $\pm$ 10.8	59.4 $\pm$ 12.9	36.6 $\pm$ 7.9	60.8 $\pm$ 12.0
0.10	37.6 $\pm$ 10.1	59.2 $\pm$ 12.0	35.0 $\pm$ 8.3	58.8 $\pm$ 11.7
0.20	30.4 $\pm$ 9.0	59.8 $\pm$ 11.6	23.2 $\pm$ 6.0	56.0 $\pm$ 10.8

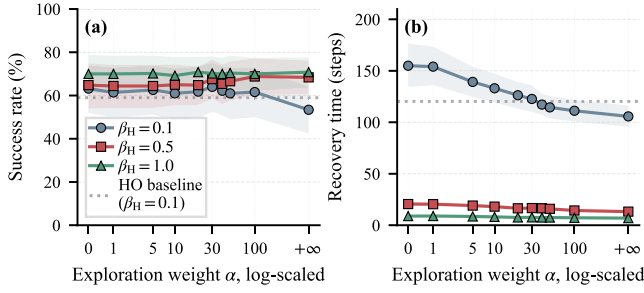


Fig. 6. Sensitivity to the exploration weight α under intention switching. (a) Success rate. (b) Recovery time, defined as the number of steps required for the inferred intention to recover after an intention switch. Curves correspond to different simulated-human rationality levels β_H . The dotted line denotes the HO baseline under $\beta_H = 0.1$.

budget, authority discretization, and the number of intention hypotheses.

Fig. 6 shows that the success rate changes only moderately over a broad range of non-extreme α values, indicating that IG3A does not rely on a narrowly tuned exploration weight. The trends also reflect the exploitation–exploration trade-off: when α is too small, the controller relies mainly on the current intention estimate and the benefit of actively seeking informative human inputs is limited; when α becomes extremely large, the pure-exploration setting ($\alpha = +\infty$) shortens recovery time but reduces success rate under the low-rationality condition. Thus, faster belief adaptation alone is insufficient for good task performance, and a moderate exploration weight is needed to balance intention exploration with control stability. Accordingly, $\alpha = 30$ is selected from this stable non-extreme range for the main experiments.

Fig. 7 evaluates robustness to rationality-model mismatch by varying $r = \hat{\beta}/\beta_H$. IG3A maintains relatively stable success rate and reward over $r \in [0.5, 2]$, indicating that the method can tolerate moderate calibration errors in $\hat{\beta}$. CB shows a larger performance drop when $\hat{\beta}$ substantially underestimates β_H , likely because it maps inferred confidence directly to authority allocation and is therefore more sensitive to distorted belief confidence. By combining the belief-dependent information term with the task-value term, IG3A is less directly tied to confidence magnitude alone. In practice, $\hat{\beta}$ can be estimated from a short calibration dataset by maximum likelihood or selected by validation performance, while extremely small or

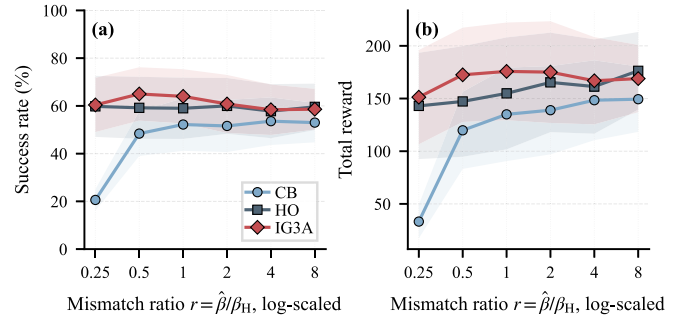


Fig. 7. Sensitivity to rationality-model mismatch. The mismatch ratio is defined as $r = \hat{\beta}/\beta_H$, where β_H is the simulated-human rationality parameter and $\hat{\beta}$ is the inference parameter. (a) Success rate. (b) Total reward.

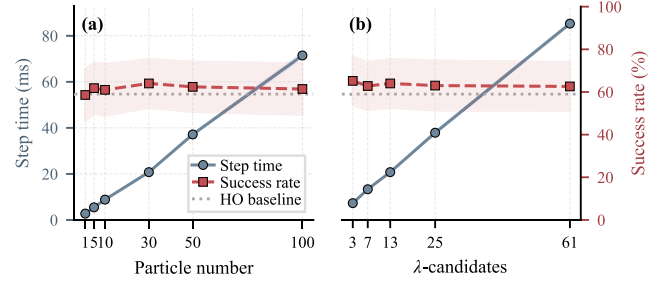


Fig. 8. Effect of sampling budget on performance and runtime. (a) Particle-number sweep. (b) λ -candidate-number sweep. The λ -candidate number denotes the number of discretized authority candidates. Each panel reports per-step computation time and success rate. The dotted line denotes the HO baseline success rate.

large values should be avoided because they may make the likelihood too flat or too overconfident.

Fig. 8 studies the sampling budget of the Monte Carlo authority evaluation, including the number of particles and the number of discretized authority candidates. The measured per-step computation time increases approximately linearly with both parameters, consistent with the complexity analysis $O(MN(C_{dyn} + K))$. In contrast, task performance varies more mildly over the tested ranges. Increasing the particle number beyond a moderate budget provides limited additional gains, and increasing the number of λ candidates has little effect on success rate in this task. These results suggest that very fine authority discretization is not necessary for the tested LunarLander setting. The main setting uses 30 particles and 13 authority candidates as a conservative non-extreme configuration: it avoids an overly coarse authority grid while keeping the per-step computation time well within the real-time range.

Taken together, the parameter sweeps provide empirical tuning guidelines for IG3A: α should be chosen from a stable non-extreme range to avoid both purely exploitative and overly exploratory behavior, while N and M should be selected to provide sufficient entropy-estimation accuracy and authority resolution under the required control frequency. Accordingly, the main experiments use $\alpha = 30$, $N = 30$ Monte Carlo particles, and $M = 13$ authority candidates.

Finally, Fig. 9 evaluates computational scalability with respect to the number of intention hypotheses K . The per-step runtime increases approximately linearly with K , consistent with the complexity analysis $O(MN(C_{dyn} + K))$ when the

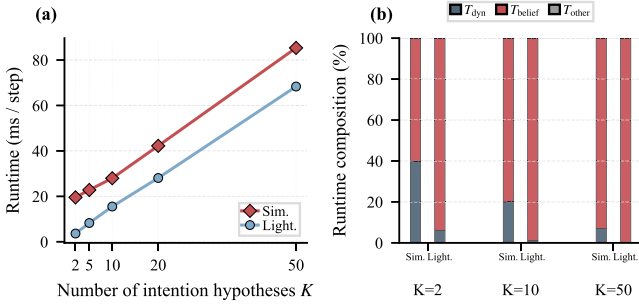


Fig. 9. Runtime scalability with respect to the number of intention hypotheses K . (a) Per-step runtime under simulator-based and lightweight prediction models. (b) Runtime composition for representative values of K , decomposed into dynamics prediction, belief update, and other computation.

numbers of authority candidates and particles are fixed. The K -dependent computation includes belief update, entropy computation, and intention-conditioned action evaluation. In Fig. 9, “Sim.” denotes simulator-based one-step prediction, whereas “Light.” denotes a lightweight linear state-space prediction model used to reduce dynamics-prediction cost and high-light intention-space computation. The runtime composition shows that dynamics prediction remains non-negligible under simulator-based prediction, while intention-space computation becomes the main cost under the lightweight model.

Under the default 2-Spread setting ($K = 2$), IG3A runs at approximately 20 ms per step, corresponding to nearly 50 Hz. Even when the number of candidate intentions increases to $K = 50$, which is much larger than the default setting and already exceeds the size of many target-based shared-control tasks, the runtime remains below 90 ms per step under simulator-based prediction, corresponding to more than 10 Hz. These results indicate that the current implementation is computationally feasible for the real-time settings considered in this paper. For larger-scale deployments, optimization should primarily target belief update, entropy computation, and intention-conditioned action evaluation; when simulator-based prediction is used, forward dynamics propagation should also be considered.

C. Human-in-the-Loop LunarLander Experiment

1) *Setup*: Human inputs were collected using an Xbox One S 1708 controller. For CB, HO, and IG3A, the same intention-inference module, autonomy controller, and authority-allocation hyperparameters as in Sec. V-B1 were used. IDA was evaluated under the same task environment, input interface, and matched trial conditions as the other methods. The rationality coefficient $\hat{\beta}$ in the likelihood model was estimated from calibration trials following the maximum-likelihood procedure in [14]. The simulator was operated at 30 Hz for all control modes.

Eighteen participants (7 female, 11 male; aged 20–31) provided informed consent and completed a 15-min training session. Each participant performed six formal trials for each method, including CB, HO, IDA, and IG3A, under the same 2-Spread goal configuration used in the main simulated-human experiment. For each method, three trials involved an assigned

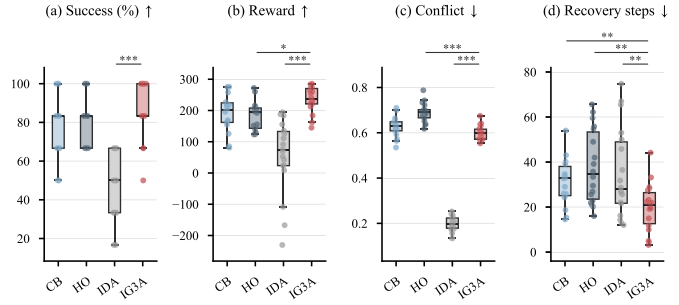


Fig. 10. HITL LunarLander results across 18 participants. Panels report (a) success rate, (b) total reward, (c) human-machine conflict, and (d) recovery steps; arrows indicate the preferred direction. Human-machine conflict is computed as the trial-averaged Euclidean distance between the participant input and the executed shared-control action. Significance markers follow Sec. V-A3.

intention switch from the left goal to the right goal, and three involved an assigned switch from the right goal to the left goal. The currently assigned goal was shown on the screen, and participants were instructed to generate control inputs toward the displayed goal, including after each switch. The assigned-goal cue was provided only to the participant; the autonomous agent inferred the intention from the observed human inputs. The same scenario seeds, switch conditions, and disturbance sequences were used across methods for each participant, and the method order was randomized to reduce learning effects.

2) *Results*: Fig. 10 reports participant-level results for the HITL LunarLander experiment. IG3A achieved the best overall task performance, obtaining the highest success rate and reward while also reducing human-machine conflict relative to CB and HO and requiring the fewest post-switch recovery steps. These results indicate that the advantage of information-aware authority allocation is not limited to simulated-human policies, but also appears when the input evidence is generated by real participants. Overall, the HITL results provide complementary evidence that IG3A can improve task execution, reduce human-machine conflict against conventional belief-based baselines, and support faster post-switch belief adaptation under changing human intentions. A supplementary mechanism analysis of belief and information-gain dynamics is provided in the Supplementary Material.

D. Physical Robot Human-in-the-Loop Experiment

The physical robot experiment evaluates IG3A in a real teleoperation loop with camera-based target detection, human inputs, and robot response delays. The hardware comparison focuses on CB, HO, and IG3A, which can be directly instantiated on the same robot inference-and-control stack with the same perception, intention-inference, and low-level control modules. The learning-based IDA baseline is included in the LunarLander experiments, but is not included here because deploying it on the robot would require a separate task-specific data-collection and training pipeline.

1) *Setup*: The physical task was designed as a multi-target robotic reaching task with intention switching. As shown in Fig. 11, the setup consisted of a 6-DoF Realman RM-65B robotic arm, a RealSense D435 depth camera, three target

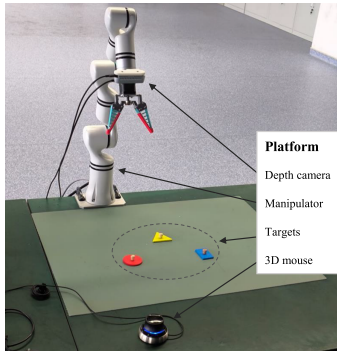


Fig. 11. Physical robot experiment setup, including the Realman RM-65B robotic arm, three target objects, RealSense D435 depth camera, and 3D SpaceMouse input device.

objects placed on a table, and a 3Dconnexion SpaceMouse Compact used as the human input device.

The robot executed incremental Cartesian commands while keeping the end-effector orientation fixed. The action was defined as $\mathbf{a} = [dx, dy, dz]^T$, with each component clipped to ± 0.03 m for both human and autonomous commands. All compared methods used the same robot interface, target detections, command limits, a 5-Hz control loop, and intention-inference module. The rationality coefficient $\hat{\beta}$ in Eq. (17) was set to 2 for all methods. For CB, the confidence-to-authority parameters were set to $\delta_1 = 0.8$, $\delta_2 = 0.2$, $\sigma_1 = 0.2$, and $\sigma_2 = 0.8$. For IG3A, λ was discretized over $[0.2, 0.8]$ with a step size of 0.05, the number of particles was set to 50, and $\alpha = 50$. Runtime profiling over 500 steps showed that the active control loop required 102.77 ± 1.23 ms per step, below the 200-ms control period of the 5-Hz loop. The IG3A arbitration itself required 31.84 ± 0.91 ms per step, indicating that IG3A computation was not the deployment bottleneck. A SpaceMouse-triggered pulse-response test over 80 trials measured a mean input-to-motion-onset latency of 194.91 ms and a mean input-to-settled time of 917.32 ms. These results indicate that the physical experiment involved non-negligible input-to-motion delay, actuation delay, and settling effects rather than ideal instantaneous control.

For the forward prediction used in the Monte Carlo evaluation, a local kinematic model was adopted:

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{a}_t + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_{\text{pred}}^2 \mathbf{I}),$$

where $\mathbf{x}_t$ is the measured end-effector position and $\mathbf{a}_t$ is the executed Cartesian command. The noise scale $\sigma_{\text{pred}} = 1.84 \times 10^{-4}$ m was estimated from empirical one-step prediction residuals collected under the same robot control loop, after excluding invalid samples caused by inverse-kinematics failures.

Fifteen participants (6 female, 9 male; aged 23–33) took part in the experiment after providing informed consent. Each participant first completed a training session to become familiar with the SpaceMouse interface and robot motion. In the formal experiment, each participant completed 18 trials, corresponding to three methods and six target-switching scenarios. The order of the three methods was counterbalanced

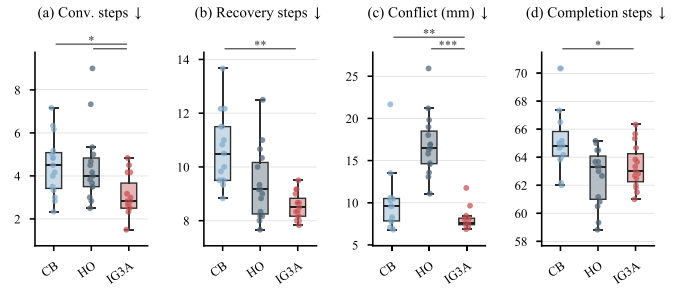


Fig. 12. Physical robot HITL results across 15 participants. Panels report (a) convergence steps, (b) recovery steps, (c) human-machine conflict, and (d) completion steps; arrows indicate the preferred direction. Human-machine conflict is computed using the same trial-averaged Euclidean-distance definition as in Fig. 10 and is reported in millimeters. Significance markers follow Sec. V-A3.

across participants, and each switching scenario was repeated once for each method.

At the beginning of each trial, the robotic arm was reset to a fixed initial configuration, and the three target positions were detected from the camera. The participant was first instructed to guide the end-effector toward an initial target. During the trial, a visual cue indicated an intention switch to a different target, and the participant was required to redirect the robot toward the new target. The autonomous agent was not given the assigned target and inferred the participant's intention only from observed human inputs. When the participant judged that the end-effector had reached the target vicinity, Button 0 on the SpaceMouse was pressed to terminate the shared-control episode.

2) *Results*: Fig. 12 reports participant-level results for the physical robot experiment. Since success was saturated in this reaching task, we report convergence steps for initial belief adaptation, recovery steps for post-switch belief adaptation, human-machine conflict as a command-level deviation metric, and completion steps for task duration.

IG3A achieved the best overall belief-adaptation performance, requiring the fewest steps for both initial convergence and post-switch recovery. Its completion time was comparable to HO and lower than CB, while its human-machine conflict was reduced compared with both CB and HO. Overall, the physical experiment provides real-hardware evidence that IG3A improves intention adaptation in a closed-loop robotic system while reducing command-level human-machine conflict compared with the tested baselines. An exploratory post-hoc analysis of participant-level rationality and the relative benefit of IG3A across both HITL experiments is provided in the Supplementary Material.

VI. DISCUSSION

The experiments lead to two main observations. First, IG3A improves shared-control performance most clearly when intention evidence is uncertain, degraded, or changing. Second, the performance gain should not be interpreted simply as a consequence of faster belief convergence, but as the result of a dynamic trade-off between active intention exploration and immediate task execution. Different from the compared baselines, this trade-off is explicitly incorporated when

selecting the authority coefficient. When the belief is initially highly uncertain, the information-gain term tends to promote informative authority choices, which is consistent with the initial-convergence results observed across the experiments. After an intention switch, however, exploration is beneficial mainly when the additional human evidence is reliable enough to offset its immediate control cost. Therefore, under severely degraded inputs, IG3A may restrain overly aggressive exploration and does not necessarily achieve the shortest recovery time, whereas increasing human goal consistency tends to make higher human authority less harmful and more informative. This pattern helps explain why IG3A can obtain better task outcomes even when it does not always produce the fastest post-switch belief recovery.

The different trends between simulated-human and HITL experiments suggest that the input-generation process matters. Although simulated-human experiments provide controlled and repeatable comparisons, their reward-rational Boltzmann commands are aligned with the likelihood model used for intention inference, which may make belief recovery effective for several belief-based methods and cannot fully substitute for real-user validation. In contrast, real participants may generate temporally correlated, delayed, corrective, or feedback-driven inputs after observing the shared-control response, and such interaction patterns are not fully represented by the idealized simulated-input model. Therefore, the HITL results provide practically important evidence that IG3A can retain its performance benefit under real-user inputs, rather than only under model-generated inputs. Given the limited participant scale, however, this observation should be interpreted as evidence of practical feasibility under real-user inputs rather than as a conclusive explanation of the underlying human-adaptation mechanism.

Several limitations and scope assumptions should be noted. First, the present validation remains limited in participant scale and physical-system complexity. The physical robot experiment provides real-hardware evidence in a target-based reaching task, but it does not yet cover more dynamic physical scenarios, larger-scale human-subject studies, or higher-dimensional systems with more complex dynamics and kinematics. The timing and scalability analyses indicate real-time feasibility under the tested settings, but they should not be interpreted as validation on arbitrary larger-scale systems. For such deployments, computational efficiency will depend mainly on the cost of one-step forward prediction and computations over the intention space, including belief update, entropy evaluation, and intention-conditioned action evaluation.

Second, IG3A is instantiated in a linear blending interface. This choice provides a simple and interpretable authority variable for isolating active intention exploration, but it does not cover more general nonlinear arbitration, probabilistic blending, or safety-filtered shared-control policies. The same information-aware principle could be extended by replacing the scalar blending coefficient with a more general arbitration or action-generation decision variable.

Third, the current framework assumes a finite and known intention set. Extending IG3A to open-set intentions is not

equivalent to simply adding an unknown label, because the belief update, goal-conditioned assistance, and information-gain evaluation all rely on a predefined hypothesis space. If the true intention is absent, the posterior may still concentrate on an incorrect known intention and produce overconfident assistance. One possible extension is to detect persistent low predictive likelihood or inconsistent belief evidence, switch to a safe goal-agnostic fallback mode, and then use perception-based candidate generation to expand the intention set before resuming belief-based authority allocation. Future work should therefore investigate unknown-intention detection, dynamic candidate generation, goal-agnostic fallback assistance, and scalable intention-space management for less structured environments.

VII. CONCLUSION

This paper presented IG3A, an information-gain-based active authority allocation framework for shared control under human bounded rationality. The proposed method treats the blending coefficient as an information-aware decision variable and selects it by jointly considering immediate task execution and expected reduction of intention uncertainty. In this way, authority allocation is not only used to combine human and autonomous commands, but also to actively shape the evidence available for intention inference.

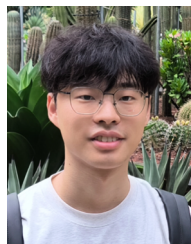
Experiments in dynamic LunarLander simulations and HITL robotic tasks showed that the proposed trade-off between active intention exploration and immediate task execution can improve overall shared-control performance when intention evidence is uncertain, degraded, or changing. The results also illustrated how exploration and exploitation play different roles across conditions: active exploration can accelerate belief adaptation when informative evidence is valuable, whereas task execution should dominate when unreliable human inputs would impose a large immediate control penalty. The HITL LunarLander and physical robot experiments further indicated that the framework can operate with real users and real hardware, and the sensitivity and timing analyses supported its robustness and real-time feasibility under the tested settings.

Future work will extend the information-aware authority allocation principle beyond linear blending to more general arbitration and safety-constrained shared-control policies. Another important direction is to relax the closed-set intention assumption by incorporating unknown-intention detection, dynamic candidate generation, and goal-agnostic fallback assistance. Larger-scale human-subject studies and more dynamic physical tasks will also be valuable for understanding how active intention exploration interacts with adaptive human behavior in practical shared-control systems.

REFERENCES

- [1] D. A. Abbink et al., "A topology of shared control systems—Finding common ground in diversity," *IEEE Trans. Hum.-Mach. Syst.*, vol. 48, no. 5, pp. 509–525, Oct. 2018.
- [2] C. Yang, Y. Zhu, and Y. Chen, "A review of human-machine cooperation in the robotics domain," *IEEE Trans. Hum.-Mach. Syst.*, vol. 52, no. 1, pp. 12–25, Feb. 2022.
- [3] S. You, Y. Kang, Y.-B. Zhao, and Q. Zhang, "Human-machine shared autonomy approach for non-full-time effective human decisions," *Sci. Sin. Inform.*, vol. 52, no. 12, pp. 2165–2177, 2022.

- [4] C. M. El Bou, F. Beck, K. D. von Ellenrieder, and S. K. Gupta, "Shared control with obstacle avoidance for UGVs," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 17109–17126, 2025.
- [5] C. M. El Bou, M. Focchi, M. Chang, M. Camurri, and K. D. von Ellenrieder, "Smooth human–robot shared control for autonomous orchard monitoring with UGVs," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 13603–13620, 2025.
- [6] J. Liang et al., "Robust shared control system for aggressive driving based on cooperative modes identification," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 11, pp. 6672–6684, Nov. 2023.
- [7] J. Liang et al., "ETS-based human–machine robust shared control design considering the network delays," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 17501–17511, 2025.
- [8] D. Zhang et al., "Human–robot shared control for surgical robot based on context-aware sim-to-real adaptation," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, Philadelphia, PA, USA, May 2022, pp. 7694–7700.
- [9] J.-A. Yow, N. P. Garg, and W. T. Ang, "Shared autonomy of a robotic manipulator for grasping under human intent uncertainty using POMDPs," *IEEE Trans. Robot.*, vol. 40, pp. 332–350, Nov. 2023, doi: [10.1109/TRO.2023.3334631](https://doi.org/10.1109/TRO.2023.3334631).
- [10] H. A. Simon, "A behavioral model of rational choice," *Quart. J. Econ.*, vol. 69, no. 1, pp. 99–118, Feb. 1955.
- [11] M. A. Goodrich and A. C. Schultz, "Human–robot interaction: A survey," *Found. Trends Hum.-Comput. Interact.*, vol. 1, no. 3, pp. 203–275, 2008.
- [12] S. Nikolaidis, Y. X. Zhu, D. Hsu, and S. Srinivasa, "Human–robot mutual adaptation in shared autonomy," in *Proc. 12th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Vienna, Austria, Mar. 2017, pp. 294–302.
- [13] J. Tan, J. Wang, S. Xue, H. Cao, H. Li, and Z. Guo, "Human–machine shared stabilization control based on safe adaptive dynamic programming with bounded rationality," *Int. J. Robust Nonlinear Control*, vol. 35, no. 11, pp. 4638–4657, Jul. 2025.
- [14] S. Jain and B. Argall, "Probabilistic human intent recognition for shared autonomy in assistive robotics," *ACM Trans. Hum.-Robot Interact.*, vol. 9, no. 1, pp. 1–23, Mar. 2020.
- [15] Q. Zhang, Y.-B. Zhao, W. Lv, and M. Chen, "Shared control with optimized arbitration for human-machine sequential decision-making," *Sci. Sin. Inform.*, vol. 53, no. 9, pp. 1768–1783, 2023.
- [16] S. Javdani, H. Admoni, S. Pellegrinelli, S. S. Srinivasa, and J. A. Bagnell, "Shared autonomy via hindsight optimization for teleoperation and teaming," *Int. J. Robot. Res.*, vol. 37, no. 7, pp. 717–742, Jun. 2018.
- [17] C. Brooks and D. Szafir, "Balanced information gathering and goal-oriented actions in shared autonomy," in *Proc. 14th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2019, pp. 85–94.
- [18] B. McMahan, Z. Peng, B. Zhou, and J. Kao, "Shared autonomy with IDA: Interventional diffusion assistance," in *Proc. Adv. Neural Inf. Process. Syst.* 37, 2024, pp. 128330–128354.
- [19] A. Jonnavittula, S. A. Mehta, and D. P. Losey, "SARI: Shared autonomy across repeated interaction," *ACM Trans. Hum.-Robot Interact.*, vol. 13, no. 2, pp. 1–36, Jun. 2024.
- [20] T. O'Neill, N. McNeese, A. Barron, and B. Schelble, "Human-autonomy teaming: A review and analysis of the empirical literature," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 64, no. 5, pp. 904–938, Aug. 2022.
- [21] H.-N. Wu, X.-Y. Jiang, and M. Wang, "Human cognitive learning in shared control via differential game with bounded rationality and incomplete information," *IEEE Trans. Artif. Intell.*, vol. 5, no. 10, pp. 5141–5152, Oct. 2024.
- [22] Y. Li, R. Cui, W. Yan, C. Feng, and S. Zhang, "Mitigating over-assistance in teleoperated mobile robots via human-centered shared autonomy: Leveraging suboptimal rationality insights," *IEEE Robot. Autom. Lett.*, vol. 10, no. 1, pp. 460–467, Jan. 2025.
- [23] U. Atan, V. R. Bharadwaj, and C. Jiang, "A probabilistic inference approach for skill-based shared autonomy in assistive robotic manipulation," *IEEE Robot. Autom. Lett.*, vol. 10, no. 9, pp. 9518–9525, Sep. 2025.
- [24] R. Luo et al., "A workload adaptive haptic shared control scheme for semi-autonomous driving," *Accident Anal. Prevention*, vol. 152, Mar. 2021, Art. no. 105968.
- [25] Z. Fang, J. Wang, Z. Wang, J. Chen, G. Yin, and H. Zhang, "Human–machine shared control for path following considering driver fatigue characteristics," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 7, pp. 7250–7264, Jul. 2024.
- [26] Z. Fang et al., "Authority allocation strategy for shared steering control considering human-machine mutual trust level," *IEEE Trans. Intell. Vehicles*, vol. 9, no. 1, pp. 2002–2015, Jan. 2024.
- [27] N. Zhong and K. Hauser, "Attentiveness map estimation for haptic teleoperation of mobile robot obstacle avoidance and approach," *IEEE Robot. Autom. Lett.*, vol. 9, no. 3, pp. 2152–2159, Mar. 2024.
- [28] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *Int. J. Robot. Res.*, vol. 32, no. 7, pp. 790–805, Jun. 2013.
- [29] A. Belsare, Z. Karimi, C. Mattson, and D. S. Brown, "Toward zero-shot user intent recognition in shared autonomy," in *Proc. 20th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2025, pp. 222–231.
- [30] H. Liu et al., "CASPER: Inferring diverse intents for assistive teleoperation with vision language models," in *Proc. Conf. Robot Learn. (CoRL)*, vol. 305, 2025, pp. 2462–2483.
- [31] D. E. Gopinath and B. D. Argall, "Active intent disambiguation for shared control robots," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 6, pp. 1497–1506, Jun. 2020.
- [32] D. E. Gopinath, A. Thompson, and B. D. Argall, "Information theoretic intent disambiguation via contextual nudges for assistive shared control," in *Algorithmic Foundations of Robotics XV*, vol. 25, S. M. LaValle, J. M. O'Kane, M. Otte, D. Sadigh, and P. Tokekar, Eds., Cham, Switzerland: Springer, 2023, pp. 239–255.
- [33] S. Reddy, A. Dragan, and S. Levine, "Shared autonomy via deep reinforcement learning," in *Proc. 14th Robot., Sci. Syst.*, Pittsburgh, PA, USA, Jun. 2018, doi: [10.15607/RSS.2018.XIV.005](https://doi.org/10.15607/RSS.2018.XIV.005).
- [34] H. Jun Jeon, D. Losey, and D. Sadigh, "Shared autonomy with learned latent actions," in *Proc. 16th Robot., Sci. Syst.*, Corvallis, OR, USA, Jul. 2020, doi: [10.15607/RSS.2020.XVI.011](https://doi.org/10.15607/RSS.2020.XVI.011).
- [35] T. Yoneda, L. Sun, G. Yang, B. Stadie, and M. Walter, "To the noise and back: Diffusion for shared autonomy," in *Proc. 19th Robot., Sci. Syst.*, Jul. 2023, doi: [10.15607/RSS.2023.XIX.014](https://doi.org/10.15607/RSS.2023.XIX.014).
- [36] C. S. Teodorescu, A. West, and B. Lennox, "Bayesian optimization with embedded stochastic functionality for enhanced robotic obstacle avoidance," *Control Eng. Pract.*, vol. 154, Jan. 2025, Art. no. 106141.
- [37] E. Yousefi, M. Chen, and I. Sharf, "Shared autonomy policy fine-tuning and alignment for robotic tasks," *Int. J. Robot. Res.*, vol. 44, no. 9, pp. 1443–1460, Aug. 2025.
- [38] C. Ezech, P. Trautman, L. Devigne, V. Bureau, M. Babel, and T. Carlson, "Probabilistic vs linear blending approaches to shared control for wheelchair driving," in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*, U.K., Jul. 2017, pp. 835–840.
- [39] K. D. von Ellenrieder, S. C. Licht, R. Belotti, and H. C. Henninger, "Shared human–robot path following control of an unmanned ground vehicle," *Mechatronics*, vol. 83, May 2022, Art. no. 102750.
- [40] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proc. 21st Int. Conf. Mach. Learn. (ICML)*, 2004, p. 1.
- [41] H. J. Jeon, S. Milli, and A. Dragan, "Reward-rational (implicit) choice: A unifying formalism for reward learning," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020, pp. 4415–4426.
- [42] A. A. Feldbaum, "Dual control theory. I," *Autom. Remote Control*, vol. 21, no. 9, pp. 874–880, 1961.
- [43] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Trans. Autom. Control*, vol. AC-19, no. 5, pp. 494–500, Oct. 1974.
- [44] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
- [45] M. Towers et al., "Gymnasium: A standard interface for reinforcement learning environments," in *Proc. Adv. Neural Inf. Process. Syst.* 38, 2025, pp. 163114–163129.



Yun-Sheng Zhao received the B.S. degree in automation from the Department of Automation, University of Science and Technology of China (USTC), Hefei, China, in 2021, where he is currently pursuing the Ph.D. degree in control science and engineering.

His research interests include AI-driven control and automation and human–machine intelligence, with a focus on shared control.



Yun-Bo Zhao (Senior Member, IEEE) received the Ph.D. degree in control theory and control engineering from the University of South Wales, Pontypridd, U.K., in 2008.

He is currently a Professor with the University of Science and Technology of China, Hefei, China, and also with the Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei. His research interests include AI-driven control and automation, human-machine intelligence, smart manufacturing, and AI-driven networked control systems.



Kangcheng Wang received the B.E. degree in automation from North China Electric Power University, Baoding, China, in 2015, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2021.

He is currently an Associate Researcher with the Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, China. His research interests include human-machine collaborative systems, intelligent manufacturing, functional testing, and modeling and optimization of complex systems.



Pengfei Li received the Ph.D. degree in control science and engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2020.

He is currently an Associate Research Fellow with the School of Information Science and Technology, USTC. His research interests include networked control systems, model predictive control, and learning based control.



Shumi Zhao received the B.S. and M.S. degrees from Hefei University of Technology in 2009 and 2012, respectively, and the Ph.D. degree from the University of Chinese Academy of Sciences in 2015.

From 2018 to 2021, he was a Post-Doctoral Researcher with The Hong Kong Polytechnic University. In 2021, he joined the Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, as a Researcher. His current research interests include bioinspired materials for advanced sensors, artificial intelligence, and embodied intelligent robots.



Cegang Li received the B.S. degree in automation from the College of Automation, Chongqing University, Chongqing, China, in 2024. He is currently pursuing the M.Eng. degree in control engineering with the Department of Automation, University of Science and Technology of China (USTC), Hefei, China.

His research interests include robotic manipulation and reinforcement learning.



Shuyue Jiang received the B.S. degree in mathematics and applied mathematics from Anhui Normal University, Wuhu, China, in 2020. She is currently pursuing the Ph.D. degree with the University of Science and Technology of China, Hefei, China.

Her current research interests include AI-driven human-machine interaction, deep reinforcement learning, embodied intelligence, and large language models for decision-making.



Yu Kang (Senior Member, IEEE) received the Ph.D. degree in control theory and control engineering from the University of Science and Technology of China, Hefei, China, in 2005.

From 2005 to 2007, he was a Post-Doctoral Fellow with the Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He is currently a Professor with Hefei University of Technology, Hefei, and also with the Department of Automation, University of Science and Technology of China. His research interests include monitoring

of vehicle emissions, adaptive/robust control, variable structure control, mobile manipulator, and Markovian jump systems.