Robust Planning with Emergent Human-like Behavior for Agents Traveling in Groups

Shih-Yun Lo[1], Elaine Schaertl Short[2], and Andrea L. Thomaz[1]
[1]the University of Texas at Austin, [2]Tufts University
yunl@utexas.edu, elaine.short@tufts.edu, athomaz@ece.utexas.edu

***Abstract*— To enable robots to smoothly interact with humans during their travels together as a group, robots need the ability to adapt their motions under environmental changes and ensure all group members' routes are feasible. To achieve this ability, robots require knowledge of the final destination and the subgoals in between. In practice, such information is seldom shared explicitly among group members, and may be frequently updated. Under this uncertain setting, maintaining travel efficiency and behavior appropriateness becomes a challenge. Previous literature approached the problem by generating compliant coordinating motions inspired by human groups, with subgoal uncertainty remaining isolated from the plan evaluation process. We show that such coordination can lead the robot to "bad" transient states where inefficient planning and lost tracking may incur. We propose to resolve the problem by formulating the coordinating motion as a Bayesian stochastic game, to plan for the robot as a group member, in the meanwhile considering the long-term effect of uncertainty during path coordination. We show that the approach improves travel efficiency and partner tracking robustness, by preventing assertive decisions during the inference update process. Moreover, the approach presents "agency", in the sense that it can generate human-like motions, which can be applied and contribute to the pedestrian simulation literature; the approach also affords variants from the human-like motions to generate robot behaviors based on sensing capabilities, contributing to the methodology of robot behavior design.**

I. Introduction

Pedestrians often travel together [1]. Travel partners share destinations or temporary subgoals. Such information is often communicated on-the-fly when local directions are to be decided. Humans can fluently adapt their motions traveling in a group, even when under frequent disturbances [2]. At decision points, e.g., intersections, we observe that there is often a leader (or more) in a group who actively decides where to go, and the follower(s), without explicit communication, can adapt their motions to quickly catch up.

The followers need to perform online subgoal inference to travel along the group, and to adapt their motions to maintain desired group shape taking into account relative position and distance to each other [1]. In robotics, this concept has been applied to enable robots with human-following capabilities, while maintaining desired group shapes [3], e.g., side-by-side walking, a configuration often seen among travelers, as shown in Fig. 1-Left. Such human-following capability has been implemented along with maximum-likelihood subgoal

*This work was supported by the US NSF (IIS-1564080, IIS-1724157) and US ONR (N000141612835, N000141612785).

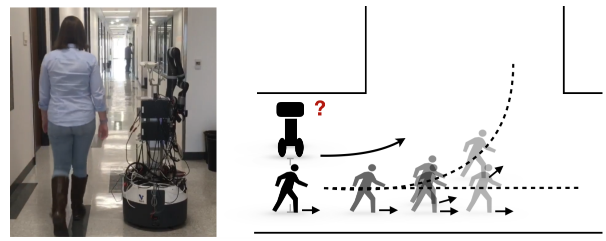

Fig. 1: Left: the robot platform operating as a travel partner. Right: the co-navigation scenario in a corridor. The robot is following the human and uncertain about where the human will be going.

estimates [4][5]. While assertively following towards the most likely subgoal showed success in relatively simple environments, such strategy, under delayed inference, can lead to bad states, e.g., with poor visibility to the other route(s) and the partner(s), or blocking partner path options. To resolve this issue, we formulate robot navigation with humans as a partial-information multi-agent planning problem, and incorporate subgoal estimation *into* the planning process for action evaluation. Under this multi-agent formulation, the robot can plan for both itself and the human travel partners, to ensure group path feasibility while maximizing the expected efficiency towards possible subgoals. We show that the proposed planner enables robust robot following when subject to inference delay, with 100% prevented lost-racking (of the travel partner) and 66% improved path efficiency. In the meanwhile, the generated robot behavior retains the compliant feature for natural co-navigation as suggested in the literature [3][4].

We noticed an emergent human-like behavioral feature resulting from this approach: hesitation – the planner would delay its action, when it is not yet certain about the subgoal but has reached states where its actions would have distinctive (and potentially bad) values under different subgoal specification. We observed such behavioral feature in human groups, and show that our approach can simulate the interactive motions during human subgroup division. As interactive agent design for small-group (or often one-to-one) interaction gains attention [6], our approach and the proposed behavioral feature contribute to the pedestrian simulation literature, building upon their approaches addressing "collective" group dynamics. The approach also affords variants of the generated behaviors based on individual sensing capabilities, which are different among humans and robotic agents.

II. Related Work

In human-robot interaction, "agency" was proposed to describe behavioral traits such as intelligence in intent-clear motions [7], collaborativeness in teamwork [4], and even cunningness in non-collaborative games [8]. Hesitation, as a behavioral trait, was studied in a human-robot object-reaching scenario [9]; by mimicking human demonstrated trajectories, robot hesitant motion was shown to be perceived as more compliant and safe by the human partner. Yet, for their approach and such motion to be applied in general human-robot interaction, it remains unclear regarding *when* and *for how long* such behavior should be exhibited to be adequate.

Individual human behavior simulation is widely studied in crowd modeling, for large-scale computer graphics applications [10] and high-density chaos analysis [11]. Due to such target applications, crowd modeling, studying both individual travelers [12] and groups [1], is often concerned with large-scale collective pedestrian behaviors. The methodology that models interaction based on "forces" among autonomous decision-making individuals is referred to as agent-based modeling [13][12]. As video games prosper, interactive agent design starts gaining attention, for which delicate realistic behaviors and small-group interactions are targeted [14][6].

For navigation, applying agent-based modeling techniques [15][16] for human-like motions enables appropriate behaviors in some situations [17][4], yet there are scenarios where incomplete decision-making models appear incapable to accommodate, e.g., inference delay can make approaches to suffer from unnatural behavior [5]. Robustness remains concerned for long-duration interaction and user engagement. Human-mimicking as a strategy to generate robot behaviors can then appear unclear of the motives for performance evaluation. Compared to those behavior-based approaches, our proposed model uses a planning formulation to evaluate action impacts, and then uses the model to generate motions with agency. For co-navigation, in this work, it is to ensure path feasibility and human-partner tracking. Such approach was also used to generate intent-communicative motions [18].

In robotics, inference mechanisms are often isolated from motion planning. Planning in human workspaces has been formulated in the single-agent setting, which fails to capture the mutual-adaptability feature in group motions [19][20]. By incorporating information uncertainties into a multi-agent planning process, our planner retains the interactive behavioral features in groups, and ensures decision robustness in the sequential decision-making process. Here we model the problem using stochastic Bayesian game, or partially-observable stochastic game [21], for their interactive policies generated based on partner/opponent modeling. Other multi-agent decision-making models, especially interactive-POMDP [22] and DEC-POMDP [23], are also suitable for our problem; planning algorithms well serve the continuous-time navigation domain, as a general control paradigm, which fits our solution concept.

III. Problem Formulation

We formulate the dynamics of group motions as a sequential optimization problem. We use a game formulation to lay out the mutual-adaptability behavioral features in the literature of crowd simulation [13][1], [2] and robot planning for human-following [3][4][5]. We then point out the issue in their behavioral assumption and propose our group follower model using stochastic Bayesian game.

A. Group Navigation: Collaborative Stochastic Games

Pedestrian groups often maintain certain *shapes* to facilitate interaction among the members, e.g., to see others' faces or stay aware of the focus or attention. Group shapes are affected by environmental conditions, such as crowd density, crowd speed, and building configuration. Under theses changes, pedestrians adapt their motions, and therefore the group shape, to coordinate with other pedestrians [1]. The shape formulation and the "mutually adaptive" feature describe the *macro* pedestrian group behaviors [1]. We therefore first model the macro pedestrian group behavior as a stochastic game. In stochastic games, N agents act at a time, here illustrated at time t: the joint-action $a_t = (a_t^1, a_t^2, ..., a_t^N) \in A$ is defined by the action spaces of all agents $A = A^1 \times A^2 ... \times A^N$; the joint-state $x_t = (x_t^1, x_t^2, ..., x_t^N) \in X$ is defined by the state spaces of all agents $X = X^1 \times X^2 ... \times X^k$. Time is discretized, and game periods are defined: at the start of each period t, each agent selects an action $a_t^i, i = 1 : N$, then the transition function $\mathcal{T} : X \times A \rightarrow X$ takes in the current state x_t and determines (probablistically) the state at the beginning of the next period x_{t+1}. The reward $\mathbf{r}_t^i$ of an agent i at time t is defined as follows: $\mathbf{r}_t^i = r^i(x_t, a_t^i, a_t^{-i}) \in \mathbb{R}$, where r^i is the agent's reward function, and $-i$ denotes all agents except i. Here we consider all agents to have the same collaborative reward function r, to maximize the group social welfare.

The optimal state-action value function Q in the single-agent setting is generally defined as: $Q(x_{t+1}) = \mathbf{r}_t(x_t, a_t) + \mathbb{E}_{x_{t+1}}[V(x_{t+1})]$. To find the optimal policy of this multi-agent problem, we first assume the agent has good modeling of the policies of other agents, π^{-i}. The optimal state-action value function of agent i while other agents use policy π^{-i} is denoted as $Q^{i|\pi^{-i}}$ and defined as:

$$Q^{i|\pi^{-i}}(x_t) = \max_{a_t^i} \mathbb{E}_{a_t^{-i}, x_{t+1}|\pi^{-i}, \mathcal{T}}[r(x_t, a_t^i, a_t^{-i}) + V^{i|\pi^{-i}}(x_{t+1})], \tag{1}$$

where $V^{i|\pi^{-i}}(x_t)$ is defined recursively as:

$$V^{i|\pi^{-i}}(x_t) = \max_{a_t^i} \mathbb{E}_{a_t^{-i}, x_{t+1}|\pi^{-i}, \mathcal{T}}[Q^{i|\pi^{-i}}(x_t, a_t^i, a_t^{-i})]. \tag{2}$$

Here the expectation $\mathbb{E}$ is taken over other agents' policies π^{-i} and the state transition $\mathcal{T}$ to accommodate noise in π^{-i}

[1]Side-by-side walking is commonly observed among groups of 2-4 members [1]. Other commonly observed configurations are "V" or "U" shapes, bent forward for 10-20 degrees, or divided side-by-side subgroups, when crowd density is high. We choose this group shape to demonstrate the proposed approach, but do not limit from applying to other shapes.

modeling and policy execution. $V^{i|\pi^{-i}}(x_t)$ is defined in the joint state space X, and dependent on the modeling of π^{-i}. The optimal action of agent i at time t is therefore:

$$a_t^{i*} = \underset{a_t^i}{\operatorname{argmax}}\, \mathbb{E}_{a_t^{-i}, x_{t+1}|\pi^{-i}, \mathcal{T}}[Q^{i|\pi^{-i}}(x_t, a_t^i, a_t^{-i})]. \quad (3)$$

B. Behavioral Assumptions in Agent-based Models

We now incorporate the goal-driven characteristics of group navigation into the individual reward function:

$$\mathbf{r}_t^{i|\theta} = r^i(x_t, a_t^i, a_t^{-i}|\theta). \quad (4)$$

The optimal policy in Eq. 3 is then amended through conditioning its terms, e.g., π^{-i}, r^i, and $Q^{i|\pi^{-i}}$, on θ,

$$a_t^{i*} = \underset{a_t^i}{\operatorname{argmax}}\, \mathbb{E}_{a_t^{-i}|\theta}[Q^i(x_t, a_t^i, a_t^{-i}|\theta)]. \quad (5)$$

This equation follows the stochastic Bayesian game formulation, in which agent rewards are parametrized by their types. Agents have partial observability to the types of other agents.

Instead of solving this stochastic Bayesian game equation based on modeling and inference of π^{-i} and θ, in previous approaches mentioned in agent-based modeling and robot navigation [2][3], local information is assumed accessible to all nearby agents, which means that θ is assumed known and shared by all agents. With the assumption that θ is of *common knowledge* to all group members, namely everyone knows they share this information and everyone knows others know that and so on [24], solving for the optimal collaborative policy for agent i in Eq. 5 converges to the solution of having a *centralized* system optimizing for the whole group,

$$a_t^{i*} = \underset{a_t^i}{\operatorname{argmax}}\, \underset{a_t^{-i}}{\max} \mathbb{E}_{x_{t+1}\sim\mathcal{T}(x_t, a_t)}[r^i(x_t, a_t^i, a_t^{-i}|\theta) + V^i(x_{t+1}|\theta)]. \quad (6)$$

This optimal policy with known θ, $\pi^*(x_t|\theta)$, is referred as the *zero-inference* policy of agent i: $\pi^{i,0}$, imposing the collective agency assumption in teammate modeling in Bayesian games [25]. It can be thought of as a policy where the leader plans for the whole group to optimize group efficiency, and the members know this plan and follow. The converged homogeneous solution $\pi^{i,0}$ describes the mutually-adaptive macro pedestrian group behavior: they reciprocally leave room for others to avoid obstacles, and expect partners to do the same. $\pi^{i,0}$ follows the formulation of $\pi^*(x_t|\theta)$, which is parametrized by θ, the *type* of an agent.

C. Proposed Follower Behavioral Model

The common knowledge assumption is however invalid for real-world applications, and can lead to inefficient motions due to false parametrization. Here we describe following behavior using the general stochastic Bayesian game formulation in Eq. 5. With group leader(s) modeled by $\pi^{-i,0}$, the follower makes observation $o_t \in O$ at time t, and use the observation history $o_{0:t-1}$ to compute the expected value over θ:

$$a_t^{i*} = \underset{a_t^i}{\operatorname{argmax}}\, \mathbb{E}_{\theta|o_{0:t-1}, \mathcal{T}}[Q^i(x_t, a_t^i, \pi^{i,0}(x_t|\theta)|\theta)]. \quad (7)$$

Here O is the observation space. Given online observations, type identification influences the optimal policy on-the-fly.

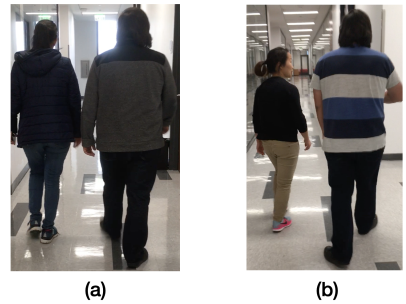

Fig. 2: (a) The follower (right) stayed slightly behind the leader (left), and (b) adapted his gait right after the leader (left) started turning.

In human groups, we found a common pattern among followers that is seldom exhibited among leaders: followers actively adjust their relative positions to stay at the rear of the group, as shown in Fig. 2. Although such behavior varies among people: they may do so throughout navigation, off-and-on, or just during some part of experience, e.g., in front of intersections [5]. This movement may seem inefficient, but it helps ensure the *future observability* of the leader motions, which is especially important when encountering distinctive route options. We associate the rear-positioning feature with information gathering for behavior generation, with sensing capability (e.g., visible range) modeled in the formulation:

$$a_{0:T}^{i*} = \underset{a_{0:T}^i}{\operatorname{argmax}}\, \mathbb{E}_{x_{0:T},\theta|\Omega^i,\mathcal{T}}[\sum_{t=0}^{T-1} r^i(x_t, a_t^i, \pi^{i,0}(x_t|\theta)|\theta) + Q_T^i(x_T, a_T^i, \pi^{i,0}(x_t|\theta)|\theta)], \quad (8)$$

where Ω^i is the observation function, affected by the agent's sensing capability. $\Omega^i : X \times A^i \rightarrow O$ samples agent i's observation given joint state x_t and agent action a_t^i. Here we consider finite-horizon lookahead T, to generate the local-observation-driven human-inspired behavior [7][12]. The multi-agent sequential optimization problem in Eq. 5 now becomes a more tractable belief planning problem in finite type space Θ. Compared to prior work in crowd modeling and robot following (formulated by Eq. 6), with Eq. 8, incorporated with agent sensing capability modeling Ω, the following agent would not assertively turn to a dead-end, go in front of the leader, or stay behind a corner obstacle, leaving the leader out of its sensing range. We here identify this uncertainty-driven observation-sensitive behavioral feature, and refer to it as having *first-order* inference for teamwork planning: $\pi^{i,1}$: here agent i needs to model other agents as having independent knowledge, which is the basic ability of *theory of mind* in human behaviors [26].

IV. Algorithm

We here detail the techniques to solve for Eq. 8. We first explain the techniques to solve for $\pi^{-i,0}$, or Eq. 3. With $\pi^{-i,0}$ solved, the goal-reaching macro group behavior can be simulated, and its performance is validated in Section V. We then solve for $\pi^{-i,1}$, or Eq. 8 with uncertainty on goal parameter θ and sensing capability modeling Ω. The planning process involves 1) search in belief space and 2) belief update

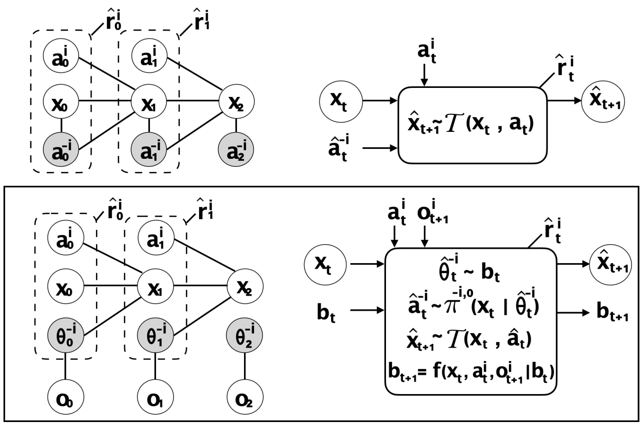

Fig. 3: Graphical model and state transition in stochastic games (Top), v.s. in stochastic Bayesian games (Bottom) for planning. In stochastic games, a_t^{-i} is directly sampled from π^{-i}; in stochastic Bayesian games, a_t^{-i} is sampled conditioned on θ, which is sampled from belief state b_t.

based on sampled observations. The procedures are illustrated in Fig. 3. The use of $\pi^{-i,0}$ for group member modeling utilizes the collective agency assumption, and restricts the strategy space for group motion sampling and for model identification, which improves the computational tractability for multi-agent planning under partial information.

A. Planning in Joint Action Space

To plan in stochastic games, it involves sampling other agents' current actions, which are hidden to the planning agent. We use heuristic search through a tree structure with finite horizon T, and use $\mathcal{T}$ for forward simulation. The tree search starts with a root node x_t. A node expands through forward-simulating the state-action pair based on sampled actions: $x_{t+1} \sim \mathcal{T}(x_t, a_t)$, and receives a reward $\mathbf{r}_t^i = r^i(x_t, a_t)$. When planning in stochastic games, both reward function $r^i(x_t, a_t^i, a_t^{-i})$ and transition function $\mathcal{T}(x_t, a_t^i, a_t^{-i})$ involve the sampling of other agents' actions a_t^{-i}, as illustrated in Fig. 3-Top. Here, with the collective agency assumption, we can solve for Eq. 6, as if all agents are controlled by agent i, for $\pi^{-i,0}$ sampling and θ identification. The search is then in the joint action space, and the worst-case complexity is: $(N|A^i|)^T$, conditioned on $\theta \in \Theta$. We apply Euclidean-distance-to-goal as heuristics in this work for value estimate at time T, $V^T(x_T|\theta)$, and use it to select which of the sampled actions to expand the search tree.

Then, applying $\pi^{-i,0}$ to solve for Eq. 8, the first action of optimal action sequence $a_t^{-i} \sim \pi^{-i,0}(x_t|\theta)$ serves as the group motion prediction for node a_t^i at state x_t.

B. Planning with Hidden Type Parameters

We apply another layer of tree search, in addition to that to solve for Eq. 6, to solve for Eq. 8. Due to the uncertainty to θ, at each time t, $\theta \in \Theta$ are sampled for $a_t^{-i} \sim \pi^{-i,0}(x_t|\theta)$, to apply state transition, as illustrated in Fig. 3-Bottom. Agent i's actions a_t^i are sampled for planning and node expansion, along which a_t^{-i} sampled for prediction, the reward and state transition are estimated conditioning on θ: $\hat{r}_t^i = \mathbb{E}_\theta r^i(x_t, a_t^i, \pi^{-i,0}(x_t|\theta)|\theta)$, and $\hat{x}_{t+1} = \mathcal{T}(x_t, a_t^i, \pi^{-i,0}(x_t|\theta))$.

When planning in partially observable environments, at the end of each time t, an observation o_{t+1} is received, based on which the beliefs of hidden variables are updated. While one would desire to sample as few observations as possible to maximize computational efficiency (especially to plan in real-time), it is important that some "key" scenarios are captured. In our navigation domain, different subgoals can lead to distinctive leader actions a_t^{-i}, which serve as observations to infer subgoal location. Here we assume a_t^{-i} are measurable at time $t+1$, serving as the observations for θ belief updates: $o_{t+1}^i \sim \Omega^i(x_t, a_t^i|\theta)$. Ω^i is affected by sensing quality and range. At each time t, we maintain a belief b_t over all possible states $\theta \in \Theta$; the belief b_t at time t can be updated through applying Bayes' rule:

$$b_{t+1}(\theta) = \eta \Omega(o_t, x_t, a_t^i|\theta) \sum_{\theta \in \Theta} b_t(\theta), \tag{9}$$

where η is a normalizing constant. Since θ is assumed static, no state transition is applied in this formulation. This model reduces the sample complexity from the observation space $|O|$ to the small belief space $|\Theta|$ and therefore improves the computational efficiency. We also apply Euclidean-distance-to-goal for expected cost-to-go estimate, $\hat{V}_T^i = \mathbb{E}_\theta[V^i(x_T|\theta)]$.

Since here we only consider agent i's own observation for o_t and use $a_t^{-i} \sim \pi^{-i,0}(x_t)$, along with a_t^i, for state transition, the planning procedure is similar to that in POMDPs: to apply state transition at state x_t based on selected action a_t^i, and to update belief based on newly received observations o_t^i, except that here we need to solve a search for every $a_t^{-i} \sim \pi^{-i,0}(x_t)$, on top of which we solve the search for $\pi^{i,1}$ (for Eq. 8). This overall search computational complexity is $(N|\Theta||A|)^T$.

C. Implementation

1) Action space: When simulating human pedestrians, we consider human states to be their positions and velocities, and sample finite actions from A^{-i} by applying constant speed change and angular velocity of the range [-1,1] m/s^2 and [-45,45] deg/s from a base velocity, which is 0.7 m/s at current walking direction. We consider the robot state as its position, body orientation, and head orientation. We use the same sample action set from A^i for robot planning.

2) Subgoal Inference: The observation o_t measures of other agents' current state x_t^{-i} for their previous actions a_{t-1}^{-i}, and can only be received when x_t^{-i} is within the visible range of agent i; if not, no Bayesian update is conducted. Based on this condition, an agent then predicts o_t based on Ω, which is a function of the previous state x_{t-1}, previous agent action a_{t-1}^i, and the hidden state θ. We here use a common velocity-based estimator for subgoal inference [27], using δ, the velocity direction (or body orientation) differences from that of the full-knowledge policy rollout $a_t^{-i} \sim \pi^{-i,0}$, conditioned on subgoal θ:

$$\Omega(o_{t+1}, x_t, a_t^i|\theta) \sim exp(-\beta\delta). \tag{10}$$

This corresponds to the intuition that an observation is sampled with higher probability when the leader takes an

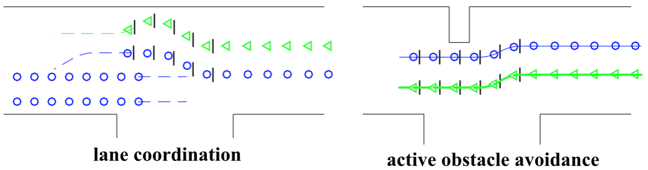

Fig. 4: Macro group behavior simulated by $\pi^{-i,0}$: the agent (marked in green triangles) plans for all agents, and therefore actively yields space for partner when encountering other groups (Left) and obstacles (Right).

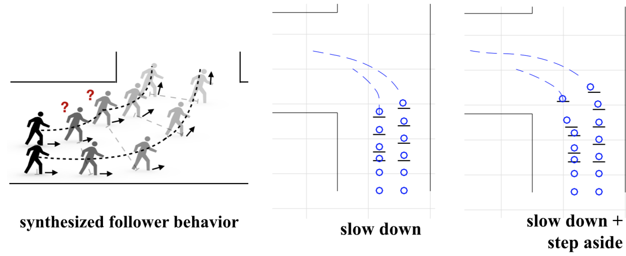

Fig. 5: Synthesized human side-by-side walking at an intersection (Left), and the simulated behaviors using $\pi^{i,1}$ (Middle and Right).

action closer that predicted by $\pi^{-i,0}$, which is conditioned on subgoal θ. $\beta > 0$, tunes how quickly the probability decays over distance. β can take in different values based on domain configuration to bias convergence rate. Here we use $\beta = 1$ for the going straight and $\beta = 1$ for turning. The probability of o_{t+1} is then used for belief update in Eq. 9.

3) Online planning: We plan in a receding-horizon fashion, to replan once once a new observation is received: at each time t, the robot executes the first planned action a_t^i, updates belief b_t based on newly received observation o_t, replans and repeats at time t+1. We consider finite horizon T for planning, covering 3 secs for follower planning $\pi^{i,1}$ with 4 secs of leader prediction using $\pi^{-i,0}$. The choice of T is based on the real-world observation that leader usually responds 3 secs before the intersection, and then the follower adapts 1 sec later.

4) Social cost function for group navigation: The planner optimizes action performance based on: travel efficiency (by time, weighted by 1), desired group configuration (quadratic regulation on shape deviation, weighted by 5), desired human walking pace (quadratic regulation on speed deviation, weighted by 5). We filter out nodes with potential collisions with the partner and the environment, detected based on a safety margin. While the desired group configuration may vary and thus needs to be estimated on-the-fly , we here omit this process, and choose the side-by-side configuration for proof of concepts.

V. Validation

We perform two studies to validate our approach: one to validate the proposed behavioral feature for pedestrian simulation, and the other for robust group following.

A. Pedestrian Group Behavior Simulation

We first validate the ability to emulate the macro group behavioral feature as proposed in the literature, and second to emulate group follower behavior observed in field study.

1) Group behavior simulation: We simulate human leader behavior based on the zero inference assumption: $\pi^{-i,0}$, and show the *macro* mutual-adaptive group behavior features (as proposed in previous literature [1][4]) in Fig. 4: in confined spaces, the planner *actively leaves space* to coordinate with an upcoming group (Fig. 4-Left), and to avoid partner collision with an upcoming obstacle (Fig. 4-Right).

With the above leader predictive model $\pi^{-i,0}$, which assumes the follower to also know the subgoal (therefore not guiding the follower explicitly), we simulate human follower behavior with $\pi^{i,1}$. The follower then plans while actively sensing the leader's subgoal, as shown in Fig. 5. We simulate with the sensing range of [-75,75] deg from head orientation, causing the simulated follower to slow down and stay slightly behind the leader once subjected to direction uncertainty. With varied safety margin, the follower may stay tight the leader (with margin=0.8m, as shown in Fig. 5-Middle), or yield space (with margin=1.1m, as shown in Fig. 5-Right) to prevent itself from blocking the leader to turn, which was suggested as preferable behavior [5]. Yet, when simulating with a robot sensing range covering [-120,120] degree, the slowing-down feature is eliminated, as shown in Fig. 7-Right.

2) Field study: to validate that our model produces realistic pedestrian behavior, we use a case study approach with real-time pedestrian data. Under exempt IRB approval, video data was collected from an overhead view, in a university atrium. One case is highlighted in Fig. 6.

At t=0 (annotated in seconds), the group of four was divided as an individual separated from the original group (location marked with blue circles); she slowed down and then changed direction from t=-4. The follower at the rear slowed down starting at t=0, changed orientation at t=2, and, qualitatively, appeared indecisive. After a short conversation, he changed walking direction at t=4. The follower (location marked with green triangles) is simulated with even initial prior at t=0 on which of the two subgroups to follow, and the hesitant behavior was simulated by delaying the belief update to turning until t=4. Before then, the original group was reformed, therefore no cost is assigned to maintain group configuration. The leader waited for the follower and they started to walk again at t=6. The dashed lines are the trajectories planned up to 3 sec ahead. The leader slowing-down motion from t=1 to t=5 follows a handed-coded speed profile to match the recorded locations.

B. Follower Performance under Inference Delay

As a robot group following strategy, we evaluate the path efficiency and partner tracking robustness when subject to inference delay, which is a common issue in robotics.

We choose cluttered environments to demonstrate the capability, where bad decisions can lead to unrecoverable states, given robot non-holonomic dynamics and limited sensing capability. We choose narrow corridor intersection for such design, with an obstacle at the corner which potentially blocks the robot's view to continuously track the partner, shown in Fig. 7. We implemented the state-of-the-art robot following approach as the baseline [5], in comparison

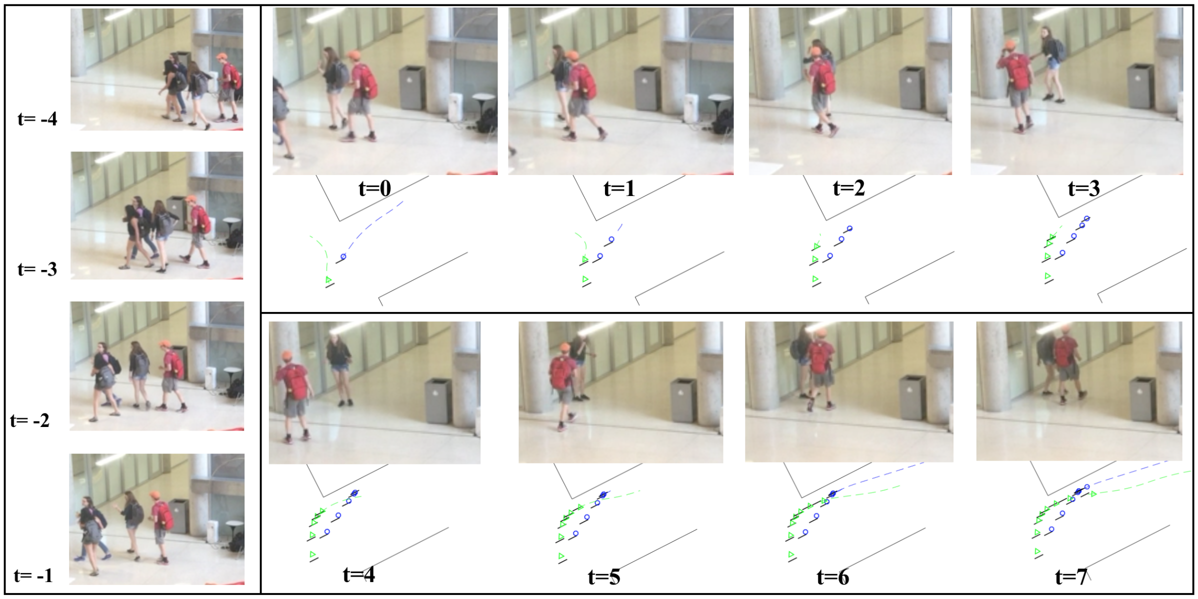

Fig. 6: Human subgroup division and its simulation by $\pi^{i,1}$: four pedestrians initially traveled as a group (t=-4); one member (marked with blue dots) gradually detached from the group when close to the intersection (t=-3 to -1). Then the one at the rear (marked with green triangle) hesitated about which to follow (t=0 to 3), later turned to catch up the waiting member (t=4 to 6), and formed a new subgroup (t=7).

with our planner under sensing range of [-120,120] deg, which is achievable with common Lidars. Experiments are conducted with 20 trials of randomized initial group locations with state estimation noise. Even priors are initialized on each local direction. Example paths can be seen in Fig. 7: as our approach is aware of future observations during the planning process, more robust motion is generated.

We report the travel time delay as the path quality measure, compared to that predicted by the zero-inference policy $\pi^{i,0}$ (that is, with a correct prior). While the baseline had an average delay of $3.53s$, our planner had an averaged delay of $1.19s$ (saving 66%). Among all trials, our planner experienced *zero* lost tracking, while the baseline experienced 10. An example is shown in Fig. 7-Middle: when belief converged at $t = 10$, the baseline had gone close to the obstacle, making it lost track of its travel partner.

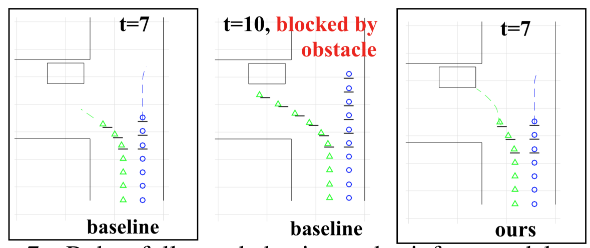

Fig. 7: Robot follower behavior under inference delay: our planner (Right) evaluates action values based on predicted future leader actions, therefore prevents bad state values (Left), which lead to future lost track of partner (Middle).

C. Discussion

1) Robustness and communication: while we showed that robust decision-making can prevent robots from losing tracking of humans and inefficient motions due to false beliefs or delayed inference, communication is a common strategy to actively resolve the negative impact of uncertainties. While robust planning helps to ensure durable deployment, communication through eye contacts, gestures, or conversations is often seen in human groups, to coordinate or to explicitly discuss about new sugoals. Partner uncertainty detection and the robot responses are therefore acknowledged as important future work to enable smooth co-navigation with humans [18].

2) Behavior design methodology: while human-emulation has been a successful methodology to generate socially- and functionally competent robot behaviors [28][29][30], our model, out of an optimization formulation, generates robot behaviors that *can be* human-like. We also show that emulating humans can lead to inefficient robot motions given sensing capability differences, e.g., robots with full-range sensing and fast inference can stay tight to the partner without concerning observability to partner motions. Our model then serves as a tool to generate robot behavior with quantitative performance improvement, in the meanwhile provides for qualitative behavioral analysis of the emergent human-like behavior.

VI. Conclusion

In this paper we present a mathematical framework for planning for group agents, implemented in group navigation, which eliminates the need for full-knowledge behavioral assumptions as used in the literature. The multi-agent partial-information planner enabled robust following performance subject to subgoal information uncertainty. When applied with limited human sensing range, it exhibited motions emulating real-world group behaviors observed in field study. Our approach generated robot behavior with improved objective performance and provided behavior-based analysis of its emergent human-like motions, serving as a tool for future human-emulating agent design.

REFERENCES

[1] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, “The walking behaviour of pedestrian social groups and its impact on crowd dynamics,” *PloS one*, vol. 5, no. 4, p. e10047, 2010.

[2] L. Huang, J. Gong, W. Li, T. Xu, S. Shen, J. Liang, Q. Feng, D. Zhang, and J. Sun, “Social force model-based group behavior simulation in virtual geographic environments,” *ISPRS International Journal of Geo-Information*, vol. 7, no. 2, p. 79, 2018.

[3] M. Kuderer and W. Burgard, “An approach to socially compliant leader following for mobile robots,” in *International Conference on Social Robotics*. Springer, 2014, pp. 239–248.

[4] L. Y. Morales Saiki, S. Satake, R. Huq, D. Glas, T. Kanda, and N. Hagita, “How do people walk side-by-side?: using a computational model of human behavior for a social robot,” in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 2012, pp. 301–308.

[5] R. Murakami, L. Y. Morales Saiki, S. Satake, T. Kanda, and H. Ishiguro, “Destination unknown: walking side-by-side without knowing the goal,” in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 2014, pp. 471–478.

[6] O. Macindoe, L. P. Kaelbling, and T. Lozano-Pérez, “Pomcop: Belief space planning for sidekicks in cooperative games.” in *AIIDE*, 2012.

[7] S.-Y. Lo, K. Yamane, and K.-i. Sugiyama, “Perception of pedestrian avoidance strategies of a self-balancing mobile robot,” in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS’19*. IEEE, 2019.

[8] E. Short, J. Hart, M. Vu, and B. Scassellati, “No fair!! an interaction with a cheating robot,” in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2010, pp. 219–226.

[9] A. Moon, C. A. Parker, E. A. Croft, and H. Van der Loos, “Design and impact of hesitation gestures during human-robot resource conflicts,” *Journal of Human-Robot Interaction*, vol. 2, no. 3, pp. 18–40, 2013.

[10] A. Treuille, S. Cooper, and Z. Popović, “Continuum crowds,” in *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3. ACM, 2006, pp. 1160–1168.

[11] N. Pelechano, J. M. Allbeck, and N. I. Badler, “Controlling individual agents in high-density crowd simulation,” in *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*. Eurographics Association, 2007, pp. 99–108.

[12] A. Turner and A. Penn, “Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment,” *Environment and planning B: Planning and Design*, vol. 29, no. 4, pp. 473–490, 2002.

[13] D. Helbing and P. Molnar, “Social force model for pedestrian dynamics,” *Physical review E*, vol. 51, no. 5, p. 4282, 1995.

[14] S. Lim and B. Reeves, “Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player,” *International Journal of Human-Computer Studies*, vol. 68, no. 1-2, pp. 57–68, 2010.

[15] F. Zanlungo, T. Ikeda, and T. Kanda, “Social force model with explicit collision prediction,” *EPL (Europhysics Letters)*, vol. 93, no. 6, p. 68005, 2011.

[16] ——, “Potential for the dynamics of pedestrians in a socially interacting group,” *Physical Review E*, vol. 89, no. 1, p. 012811, 2014.

[17] M. Shiomi, F. Zanlungo, K. Hayashi, and T. Kanda, “Towards a socially acceptable collision avoidance for a mobile robot navigating among pedestrians using a pedestrian model,” *International Journal of Social Robotics*, vol. 6, no. 3, pp. 443–455, 2014.

[18] S.-Y. Lo, E. S. Short, and A. L. Thomaz, “Planning with partner uncertainty modeling for efficient information revealing in teamwork,” in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 319–327.

[19] B. D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. Srinivasa, “Planning-based prediction for pedestrians,” in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, 2009, pp. 3931–3936.

[20] T. Kruse, P. Basili, S. Glasauer, and A. Kirsch, “Legible robot navigation in the proximity of moving humans,” in *Advanced Robotics and its Social Impacts (ARSO), 2012 IEEE Workshop on*. IEEE, 2012, pp. 83–88.

[21] E. A. Hansen, D. S. Bernstein, and S. Zilberstein, “Dynamic programming for partially observable stochastic games,” in *AAAI*, vol. 4, 2004, pp. 709–715.

[22] P. J. Gmytrasiewicz and P. Doshi, “A framework for sequential planning in multi-agent settings,” *Journal of Artificial Intelligence Research*, vol. 24, pp. 49–79, 2005.

[23] D. S. Bernstein, R. Givan, N. Immerman, and S. Zilberstein, “The complexity of decentralized control of markov decision processes,” *Mathematics of operations research*, vol. 27, no. 4, pp. 819–840, 2002.

[24] R. J. Aumann, “Agreeing to disagree,” *The annals of statistics*, pp. 1236–1239, 1976.

[25] M. Bacharach, “Interactive team reasoning: A contribution to the theory of co-operation,” *Research in economics*, vol. 53, no. 2, pp. 117–147, 1999.

[26] T. Hellström and S. Bensch, “Understandable robots-what, why, and how,” *Paladyn, Journal of Behavioral Robotics*, vol. 9, no. 1, pp. 110–123, 2018.

[27] V. V. Unhelkar, C. Pérez-D’Arpino, L. Stirling, and J. A. Shah, “Human-robot co-navigation using anticipatory indicators of human walking motion,” in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 6183–6190.

[28] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, “How to approach humans?: strategies for social robots to initiate interaction,” in *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM, 2009, pp. 109–116.

[29] C. Shi, M. Shimada, T. Kanda, H. Ishiguro, and N. Hagita, “Spatial formation model for initiating conversation,” *Proceedings of robotics: Science and systems VII*, pp. 305–313, 2011.

[30] K. Mizumaru, S. Satake, T. Kanda, and T. Ono, “Stop doing it! approaching strategy for a robot to admonish pedestrians,” in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 449–457.