Towards Safe and Predictable Social Navigation for Autonomous Ground Vehicles

Kazuki Mizuta
University of Washington

Karen Leung
University of Washington

Abstract

Planning safe and effective robot behavior in dynamic, human-centric environments remains a core challenge due to the need to handle uncertainty, adapt in real-time, and ensure safety. In this work, we introduce a unified generation-refinement framework bridging learning and optimization with a novel **reward-guided conditional flow matching** (CFM) model and model predictive path integral (MPPI) control. Our key innovation is in the incorporation of a **bidirectional information exchange**: samples from a reward-guided CFM model provide informed priors for MPPI refinement, while the optimal trajectory from MPPI warm-starts the next CFM generation. Using autonomous social navigation as a motivating application, we demonstrate that our approach can flexibly adapt to dynamic environments to satisfy safety requirements in real-time.

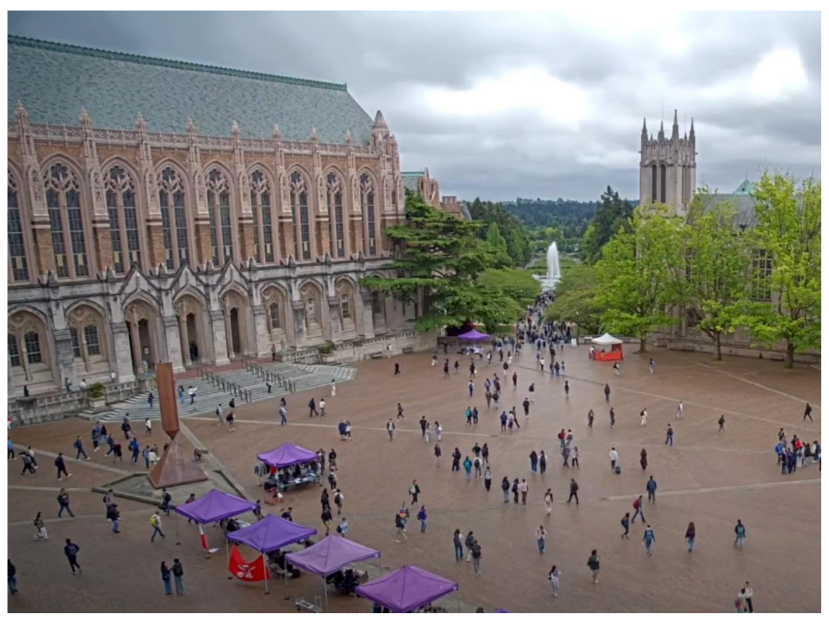


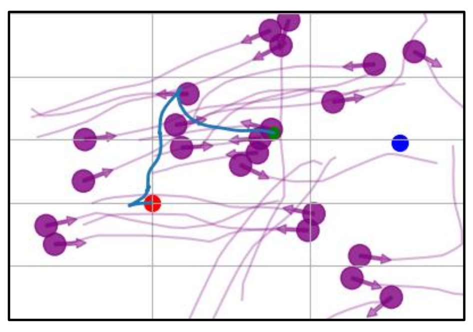
Fig. 1: Dynamic human-dense environments.

Background and Motivation

Limitations of Existing Methods:

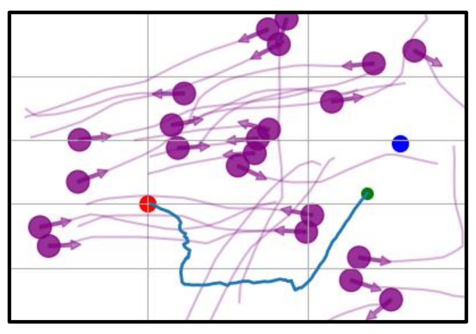
Optimization-based methods:

- ✓ Real-time planning
- ✓ Constraint handling
- ✓ Computational efficiency
- × Dynamic settings
- × Oversimplified initialization



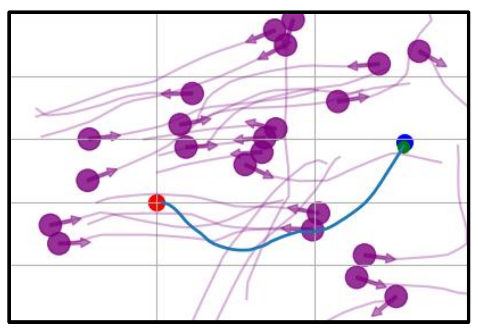
Learning-based methods (e.g., diffusion policies):

- ✓ Flexible behavior
- ✓ Multimodal uncertainty handling
- × Slow inference
- × No guarantees



Promising Direction:
Unified generation-refinement framework with conditional flow matching (CFM)

- ✓ Real-time planning
- ✓ Flexible behavior
- ✓ Multimodal uncertainty handling
- ✓ Constraint handling



Reward-Guided CFM

Key Advantage

- Enables real-time adaptation to dynamic environments without retraining
- Achieves optimal balance between generation quality and computational efficiency

Core Concept

- Incorporate reward-based guidance into the CFM trajectory generation process
- Evaluate rewards on estimated noise-free trajectories

Reward

- **Safety Reward using control barrier function (CBF):**

$$r_{\mathrm{CBF}}(\mathbf{x}_t, \mathbf{u}_t) = \gamma_t \cdot \min\{0, \dot{h}(\mathbf{x}_t) + \alpha(h(\mathbf{x}_t))\}$$

- **Goal Reward**

$$r_{\mathrm{goal}}(\mathbf{x}_T, \mathbf{x}_g) = -\|\mathbf{x}_T - \mathbf{x}_g\|^2$$

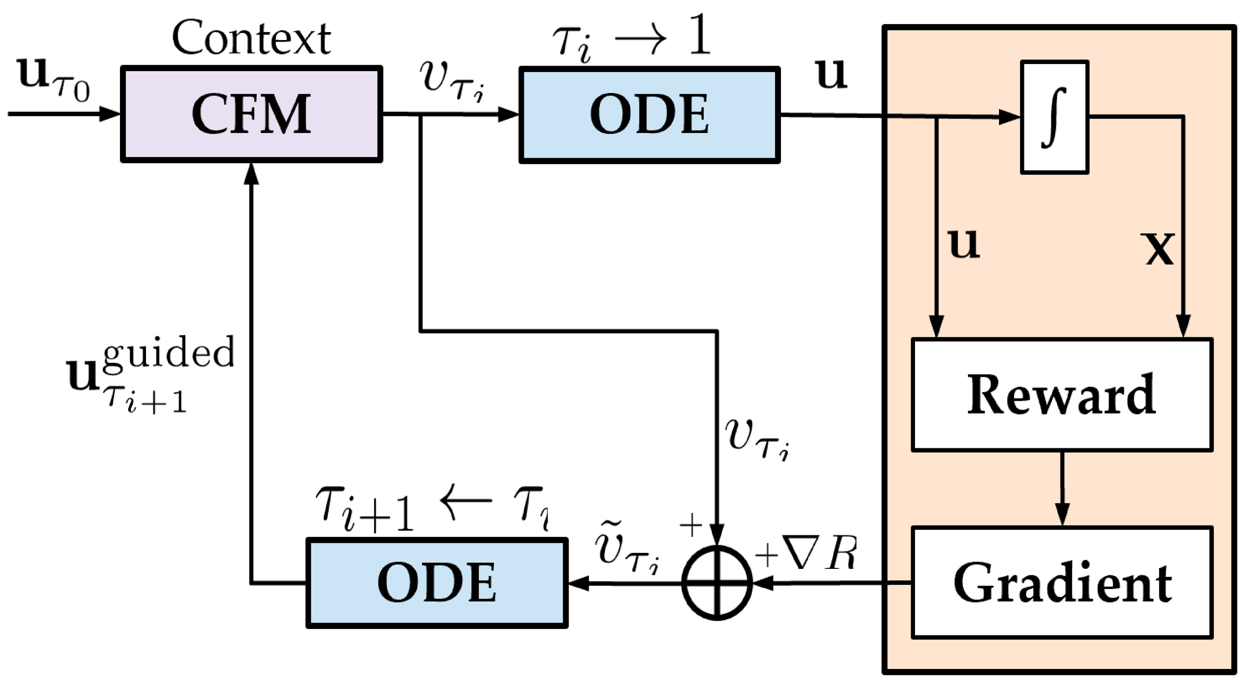


Fig. 2: Overview of the guided CFM algorithm.

Integration with MPPI

Integrate CFM framework with Model Predictive Path Integral (MPPI) control, which is compatible with general sampling-based MPC techniques

Benefits

- Uses multiple trajectories from CFM as informed samples instead of random perturbations
- Incorporates additional constraints and dynamics not covered by CFM
- Creates bidirectional feedback loop: MPC output warm-starts next CFM step

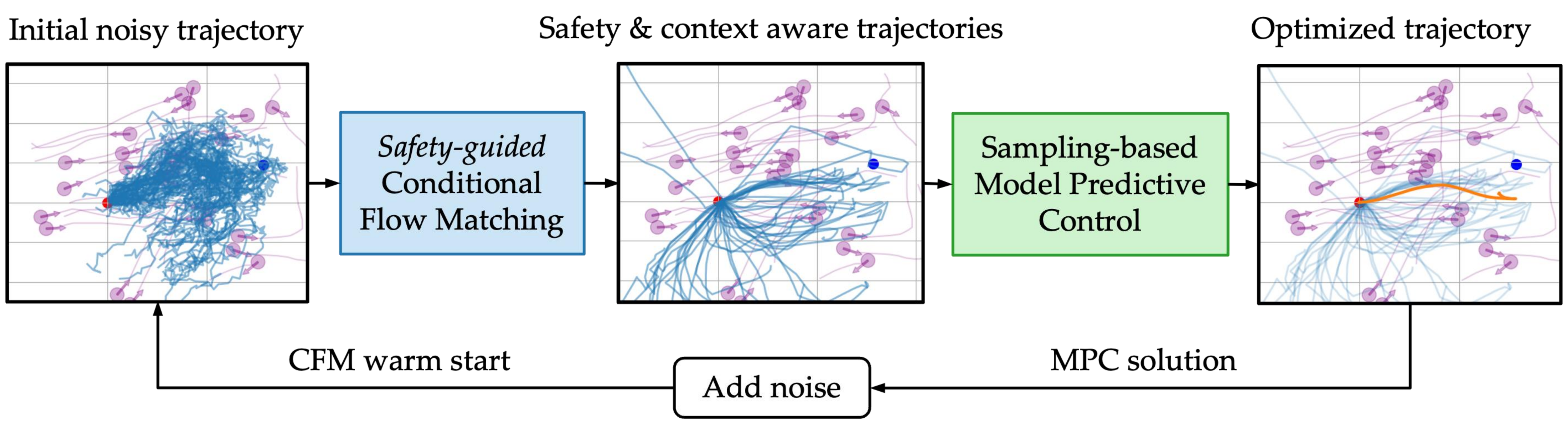


Fig. 3: Overview of the proposed unified planning framework for dynamic environments: A safety-guided CFM model generates diverse trajectories as priors for sampling-based MPC, which in turn warm-starts the next CFM sampling step.

Experiments

Baseline method:

- **MPPI:** MPPI framework utilizing a Gaussian prior
- **CFM:** A Sample from the reward-guided CFM
- **Diff-MPPI:** Guided-Diffusion model as a prior

Environments:

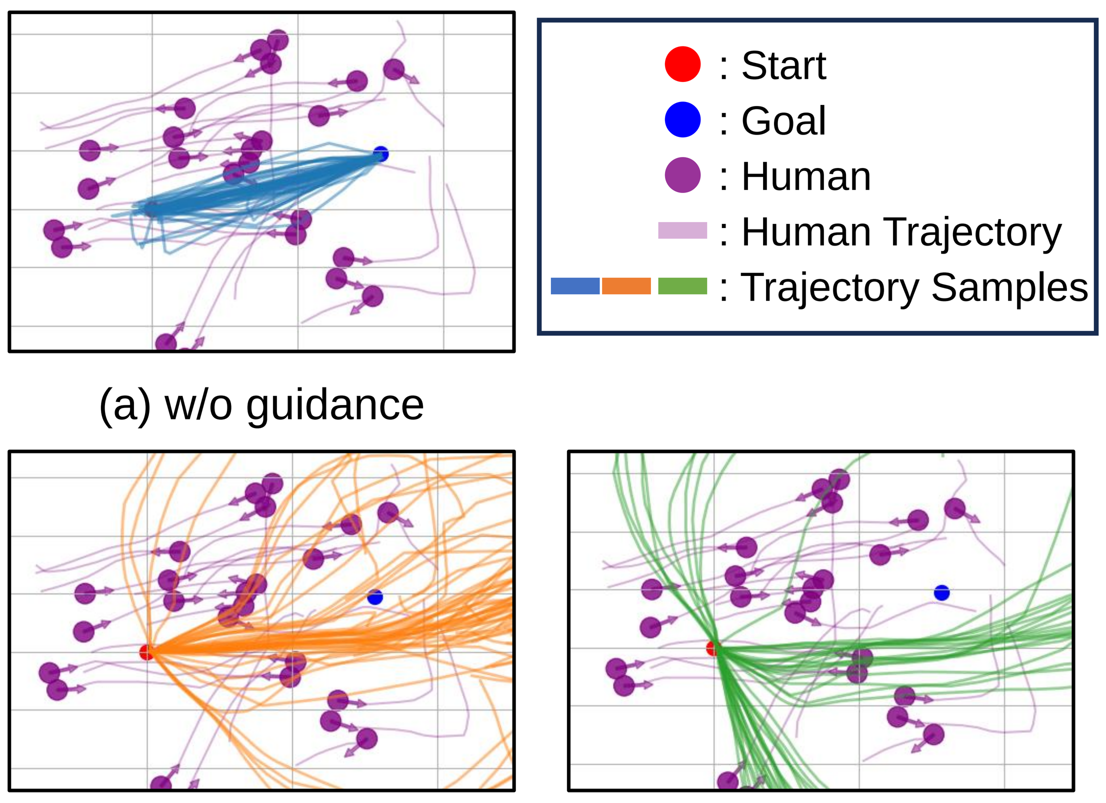
- **UCY:** pedestrians
- **SDD:** heterogenous obstacles
- **Crowd Simulation:** dense and reactive obstacles

Evaluation metrics:

- **Collision:** percentage of simulations that violate the collision radius
- **Reaching:** distance between the goal and the final state
- **Acceleration:** average linear and angular acceleration
- **Time:** computation time required to compute a control input at the current time step

Table 1: Quantitative performance comparison.

Method	Coll. (%) ↓	Reach (m) ↓	Accel. (m/s^2) ↓	Accel. (rad/s^2) ↓	Time (s) ↓
UCY Dataset					
MPPI	0.67	0.99 ± 1.79	5.09 ± 0.05	6.83 ± 0.04	**0.006**
CFM	5.67	1.06 ± 1.01	8.22 ± 3.14	7.32 ± 1.72	0.036
Diff-MPPI	0.33	0.61 ± 0.20	12.99 ± 0.66	30.16 ± 2.50	0.121
CFM-MPPI	**0.00**	$\mathbf{0.10 \pm 0.03}$	$\mathbf{1.58 \pm 0.05}$	$\mathbf{3.89 \pm 0.50}$	0.076
SDD Dataset					
MPPI	1.67	0.56 ± 0.25	5.05 ± 0.05	6.80 ± 0.04	**0.006**
CFM	5.00	1.19 ± 1.62	7.43 ± 2.58	7.64 ± 2.67	0.035
Diff-MPPI	0.67	0.62 ± 0.19	13.34 ± 0.56	31.45 ± 2.73	0.120
CFM-MPPI	**0.00**	$\mathbf{0.10 \pm 0.04}$	$\mathbf{1.54 \pm 0.05}$	$\mathbf{3.55 \pm 0.34}$	0.076
Simulated Crowd Environment with Social Force Model					
MPPI	6.67	0.82 ± 0.64	5.46 ± 0.04	6.80 ± 0.04	**0.006**
CFM	22.33	2.73 ± 2.40	10.15 ± 3.27	9.57 ± 2.41	0.036
Diff-MPPI	2.00	2.16 ± 1.97	13.74 ± 0.52	30.26 ± 2.91	0.122
CFM-MPPI	**1.00**	$\mathbf{0.49 \pm 0.69}$	$\mathbf{2.57 \pm 0.06}$	$\mathbf{4.97 \pm 0.21}$	0.076

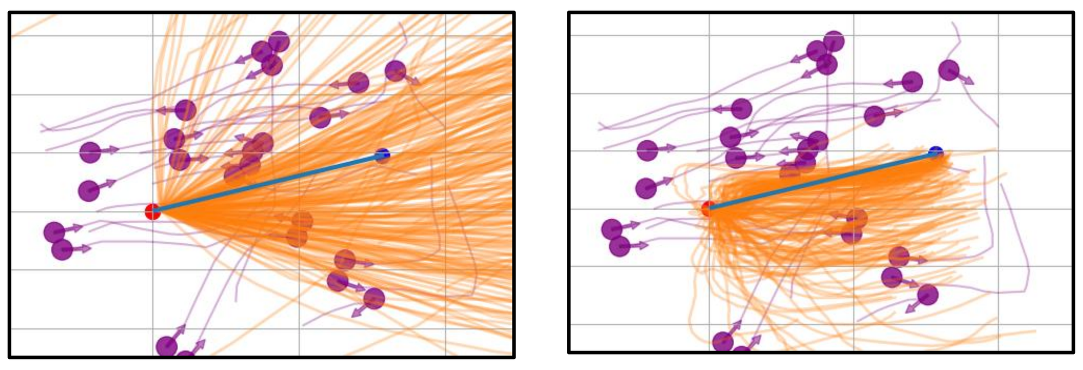


(a) w/o guidance

(b) w/ week guidance

(c) w/ strong guidance

Fig. 4: Multimodality of samples from CFM.



(a) w/o warm-start

(b) w/ warm-start

Fig. 5: Warm-Start of CFM-MPPI.

Takeaways

Safety & Efficiency

- Best performance on safety, reaching, acceleration

Real-Time Performance

- Compatible with 10Hz planning

Robustness

- Handles reactive agents
- Generates multiple trajectory candidates rather than relying on a single solution
- MPPI algorithm filters extreme behaviors for balanced navigation