We thank the reviewers for valuable feedback and will make the suggested changes. We've included additional experiments to address the existing concerns/areas of improvements. Reviewer 2: sec. C provides additional results on non-vehicle classes (i.e. bike & ped.) Reviewer 3: here we compare MFP directly to PRECOG[A] on their released CARLA data in Tab. 1. MFP *significantly* outperforms previous SOTA in [A] for 5 agent joint predictions. We also quantitatively evaluated *hypothetical* inference in Tab. 2. We report new results using the minMSD sample metric. Reviewer 5: in sec. B, we created a CARLA-based RL env. and task and proposes a simple MPF-based shooting policy (a form of MPC). We compared it with several SOTA model-free methods, demonstrating faster training and leading to a safer or more robust policy. Reviewer 6: We will release code in the near future and make the suggested clarifications. Please see detailed responses below.

REV2: The learning algorithm box was included in the Appendix. The E-step computes the true posterior distribution (step 21 of Alg. 1). MFP is a general framework and does not assume any vehicle specific dynamics or priors. MFP can perform multi-agent joint predictions of any objects, not just vehicles. (e.g. N-body problem or physics-based interactions). Sec. C. shows new results on bike and ped. prediction tasks.

REV3: We will clarify Tab.1 in the paper and add derivations of Eqs. 5, 6. DESIRE[21] is variational. Variational learning (e.g. EM or VAEs) is principled with certain guarantees on the ELBO. "inter." means interactive and hypothetical means the ability to perform conditional inference by fixing a particular agent's future trajectory. NLL is computed in closed form and is normalized by the num of timesteps and agents. For NGSIM the coordinate is in feet. Ablative studies to test how latent modes change could be performed by removing certain agents from the scene. NGSIM results did not use visual context, as visual context (grey lines of Fig.5) did not improve performance in a significant way. MFP of Tab. 1 below use 100x100x4 LIDAR rasterization as visual context. Our CARLA trajectories are 5 seconds at 20 Hz. The KL term (L171) differs from cross-entropy by a constant term (not dependent on θ_Z), so we can use them interchangeably. We compare with [A]'s CARLA dataset and MFP achieves new ***state-of-the-art*** on the most challenging 5-agents joint prediction task in Town02. [A] is closely related to MFP but one difference is that MFP can handle arbitrary number of agents while [A] (if we're not mistaken) requires a fixed num of N agents. We will certainly cite and compare/contrast with [A], and we thank the authors of [A] for providing their dataset.

REV5: We thank the reviewer for a thoughtful review and one of the original motivations was better decision making. We will add discussions to the referred RL papers. MFP can be used to learn better $p(s'|s,a)$ for model-based RL. In sec. B, we connect predictions to RL by creating a hard self-driving RL task in CARLA. We use MFP to learn good predictive models and then our policy (a form of MPC) can use Shooting methods[H] to check for future collisions within τ meters. We show episodic reward curves and also test for robustness/safety by changing the distribution of initial conditions of other agents. We achieve superior performance compare to SOTA model-free Deep RL methods.

REV6: The mentioned "Multi-modal ..." paper is a previous version of the CSLSTM paper [8] which we cite and compare with. PoV normalization rotates and translate the observations of other agents to an ego-centric frame and helps learning. Z variables are enumerated during the E-step. NLL is simply neg. log-likelihood and can be computed in closed form. MFP-1 is better than CSLSTM due our dynamic attention mechanism. MFP-1 is just a baseline to compare to other unimodal methods. Hypothetical refers to Sec 3.2 and variational learning is desirable as it is probabilistically sound. We will clarify these points and open source our code in the near future.

A. COMPARISON TO PRECOG [A] (CARLA). We train MFP (with and without LIDAR; 3, 5, and 7 modes) on the PRECOG CARLA dataset [A]. MFP is trained on 60,701 Town01 sequences for 300K updates. We report apples-to-apples comparison using minMSD metric $\hat{m}_{K=12}$ on Town02 testset for all 5 agents *jointly*. MFP (green) achieve SOTA results in Tab. 1. A quantitative eval of sec. 3.2. is in Tab. 2.

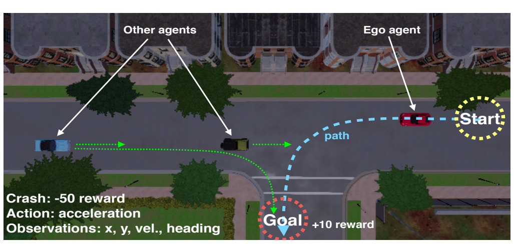
Table 1: *CARLA (PRECOG) Town02*. minMSD computed exactly as Eq. 13 of [A].

minMSD (meters)	DESIRE [21]	SocialGAN [C]	R2P2-MA [D]	ESP[A] no LIDAR	MFP5 no LIDAR	ESP[A]	MFP3	MFP5	MFP7
5 agents *joint* $\hat{m}_{K=12}$	2.422 ±0.017	1.141 ±0.015	0.770 ±0.008	1.102 ±0.011	0.842 ±0.025	0.675 ±0.007	0.641 ±0.018	0.553 ±0.013	**0.496 ±0.011**

Table 2: *Hypothetical Rollouts*. Ex. from Fig. 4(a).

$\hat{m}_{K=10}$ (meters)	MFP3 Veh1(blue)	MFP3 Veh2(green)	MFP3+Hypothetical Veh1(blue)	MFP3+Hypothetical Veh2(green)
minMSD	2.081 ± 0.25	2.765 ± 0.18	$\mathbf{1.764 \pm 0.13}$	$\mathbf{2.199 \pm 0.14}$
minFDE	3.137 ± 0.18	3.419 ± 0.31	$\mathbf{2.732 \pm 0.12}$	$\mathbf{2.742 \pm 0.26}$

B. CARLA RL ENVIRONMENT - UNPROTECTED LEFT TURN. We create an unprotected left turn task in CARLA Town05, where the objective is for Ego to safely complete an unprotected (no traffic lights) turn. Two oncoming vehicles have random initial speeds. MFP can be used in model-based RL in multiple ways: first is similar to Dyna-Q[I], where a MFP can be used to generate imagination rollouts to be added to the experience buffer. The second is an online planning algorithm (Shooting), where Ego's future action sequences are optimized to maximize for the planning reward under the learned MFP dynamics model. We compare this with several SOTA model-free methods and show that MFP-Shooting requires less sample complexity and is more robust to variations in test environment parameters.



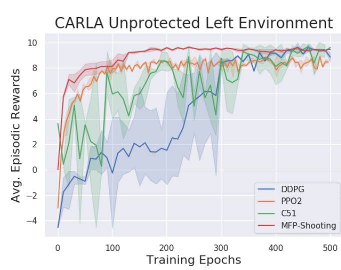


Table 3: *Testing crash rates per 100 trials. Test env. modifies the velocity & acceleration of other vehicles to test for generalization.*

Δ Env. Params	DDPG[E]	PPO2[F]	C51[G]	MFP-S τ=5m	MFP-S τ=10m
$vel: +0m/s$	2%	1%	0%	0%	0%
$vel: +5m/s$	7%	4%	5%	1%	0%
$vel: +10m/s$	13%	5%	7%	0%	0%
$acc: +1m/s^2$	9%	3%	2%	1%	0%

C. BIKE AND PEDESTRIAN PREDICTIONS. We perform additional experiments on the Stanford Drone Dataset (SDD) [J] for ped. and bike predictions. We train MFP on videos 0,1,2,4,5 of the *deathCircle* scene and test on video3. Red and blue lines are the mean predicted trajectory of two modes and the green is the predicted multi-modal log-probability density. MFP performs significantly better than baselines.

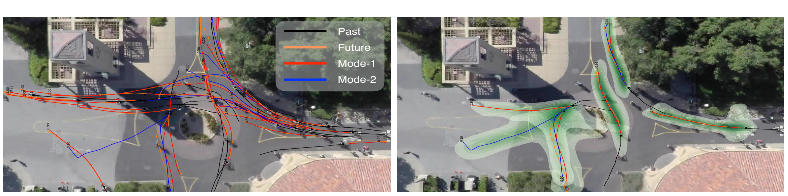


Figure 1: *Left: predicted bike trajs. Right: selected future prob. density for bikes.*

Table 4: *SDD: bike and ped. predictions on 'deathCircle' scene, video 3. Past:3 secs. Future:5 secs.* $\hat{m}_{K=12}$.

	metric	Cons Vel.	RNN	CSLSTM	MFP3	MFP5
Bike	minMSD(pixels)	10.31	8.75	8.44	5.34	4.77
Bike	Neg. LL(nats)	-	5.67	5.25	2.03	1.74
Ped.	minMSD(pixels)	4.33	3.28	3.01	2.61	2.14
Ped.	Neg. LL(nats)	-	3.39	3.07	1.44	1.31

References: **[A]** *PRECOG, Rhinehart et al. ICCV '19*. **[C]** *SocialGAN, Gupta et al. CVPR '18*. **[D]** *R2P2, Rhinehart et al. ECCV '18*. **[E]** *Deep DPG, Lillicrap et al. '15*. **[F]** *Proximal Policy Optimization, Schulman et al. '18. https://github.com/openai/baselines*. **[G]** *Dopamine: Castro et al. '18. https://github.com/google/dopamine*. **[H]** *Robust Constrained MPC, Richards A. G. Phd Thesis, '05*. **[I]** *Dyna an Integrated Architecture, Sutton '91*. **[J]** *Stanford Drone Dataset, Robicquet et al., ECCV '16*.