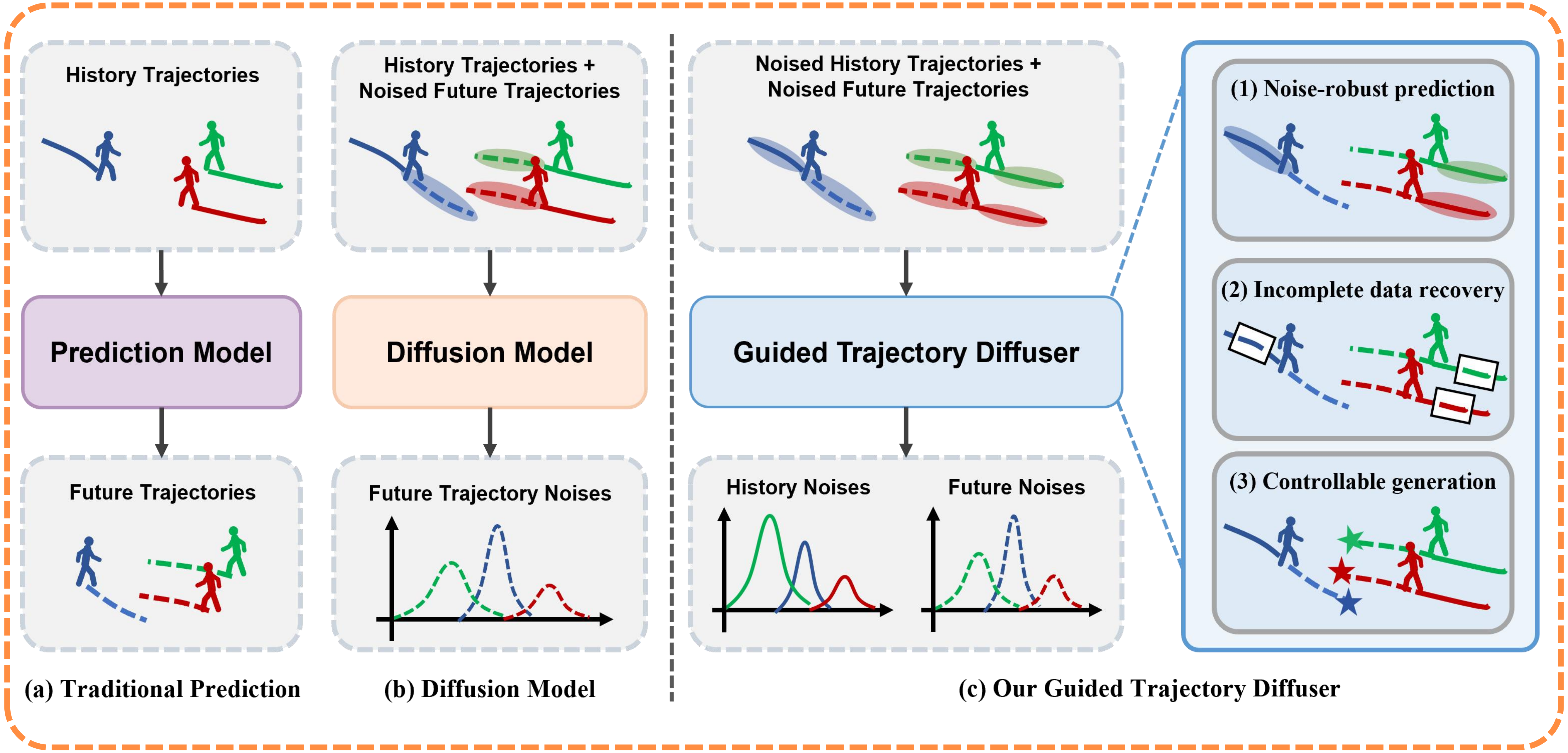
IROS '24

ABU DHABI

Haotian Lin[1,2], Yixiao Wang[1], Mingxiao Huo[3], Chensheng Peng[1], Zhiyuan Liu[3], Masayoshi Tomizuka[1]

1 University of California Berkeley 2 Tsinghua University 3 Carnegie Mellon University



IROS '24
ABU DHABI

Problem Settings

- Multi-agent trajectory prediction task is to use the observed history trajectories to jointly generate future trajectories of all the agents in the scene.

- Instead of modelling conditional distribution $p(y|E,c)$ we consider model $p(x|E)$

  $$x^i = \hat{c}^i \cup y^i = \{P^i_{-T_{his}}, \ldots, P^i_{-1}, P^i_1, \ldots, P^i_{T_{fut}}\}$$

- sample from conditional information by incorporating guidance during inference.

$$\nabla_x \log p_t(x|E,c) = \nabla_x \log p_t(x|E) + \nabla_x \log p_t(c|x,E)$$

Algorithm 1 Guided Full Trajectory Diffuser (GFTD)

Input: $\mathscr{L}(\cdot,\cdot)$, $\phi(\cdot)$, c, $\{\bar{\alpha}_t, \beta_t, \lambda_t\}_{t=0}^{T-1}$,

$x_T \sim \mathcal{N}(0,\Sigma)$

for $t = T-1, \ldots, 1$ **do**

    $\varepsilon \sim \mathcal{N}(0,\Sigma)$

    $x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\varepsilon_\theta(x_t,t)\right) + \sqrt{\beta_t}\varepsilon$

    $\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(x_t - \sqrt{1-\bar{\alpha}_t}\varepsilon_\theta(x_t,t)\right)$

    $g = -\nabla_{x_t}\mathscr{L}(\hat{x}_0, c)$

    $x_{t-1} = x_{t-1} + \lambda_t g$

    if RePaint **and** $t > 0$ **then**

        $\varepsilon' \sim \mathcal{N}(0,\Sigma)$

        $c_{t-1} = \sqrt{1-\beta_{t-1}}c_0 + \sqrt{\beta}_{t-1}\varepsilon$

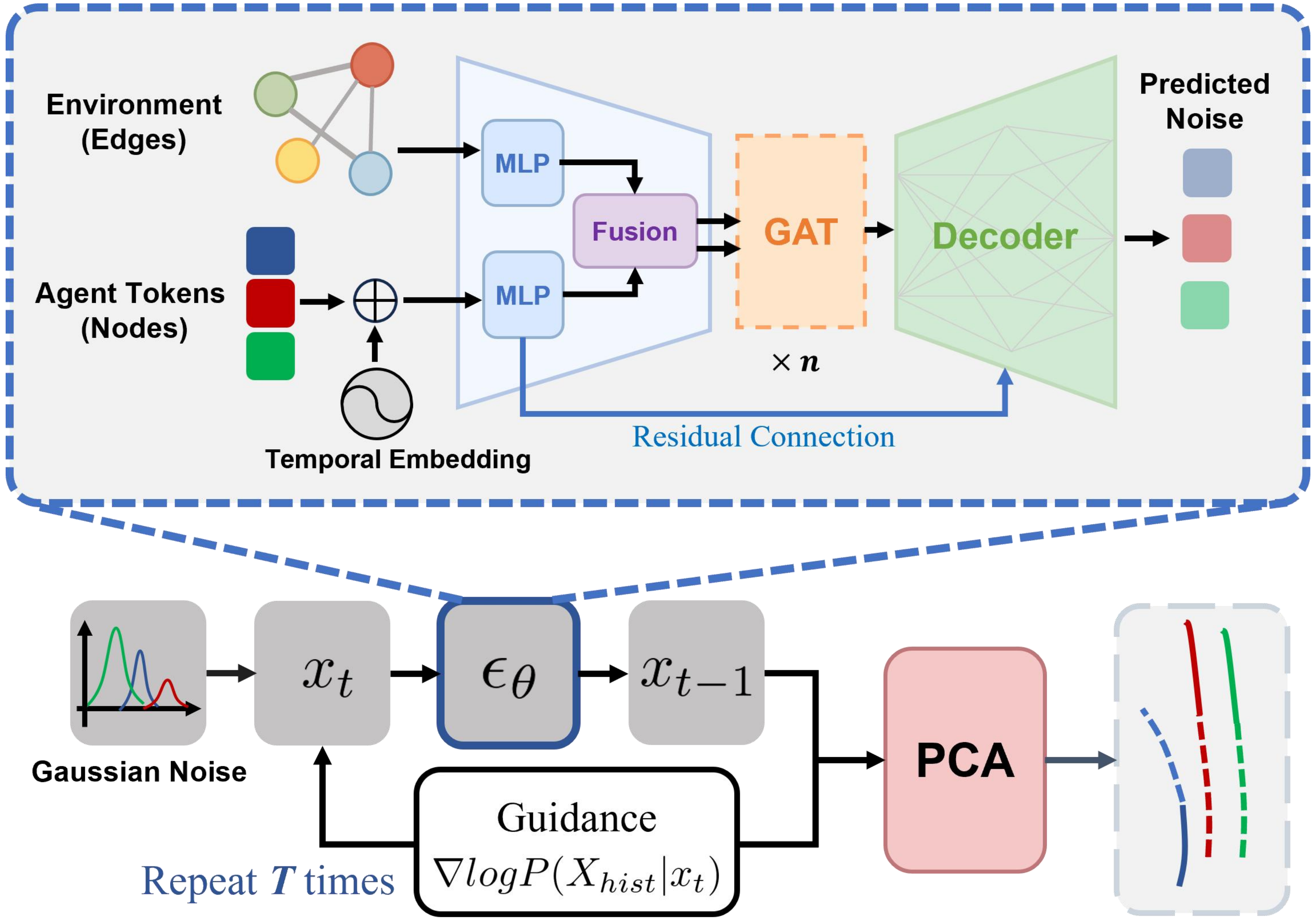
        $\phi_{his}(x_{t-1}) = c_{t-1}$

    end if

end for

$y = \phi_{fut}(x_0)$

Output: y



$$\nabla_x \log p_t(y|x) = -\lambda \nabla_{x_t} \|y - \phi(\hat{x}_0(x_t))\|_2^2 = -\lambda \nabla_{x_t} \mathcal{L}$$

Algorithm 1 Guided Full Trajectory Diffuser (GFTD)

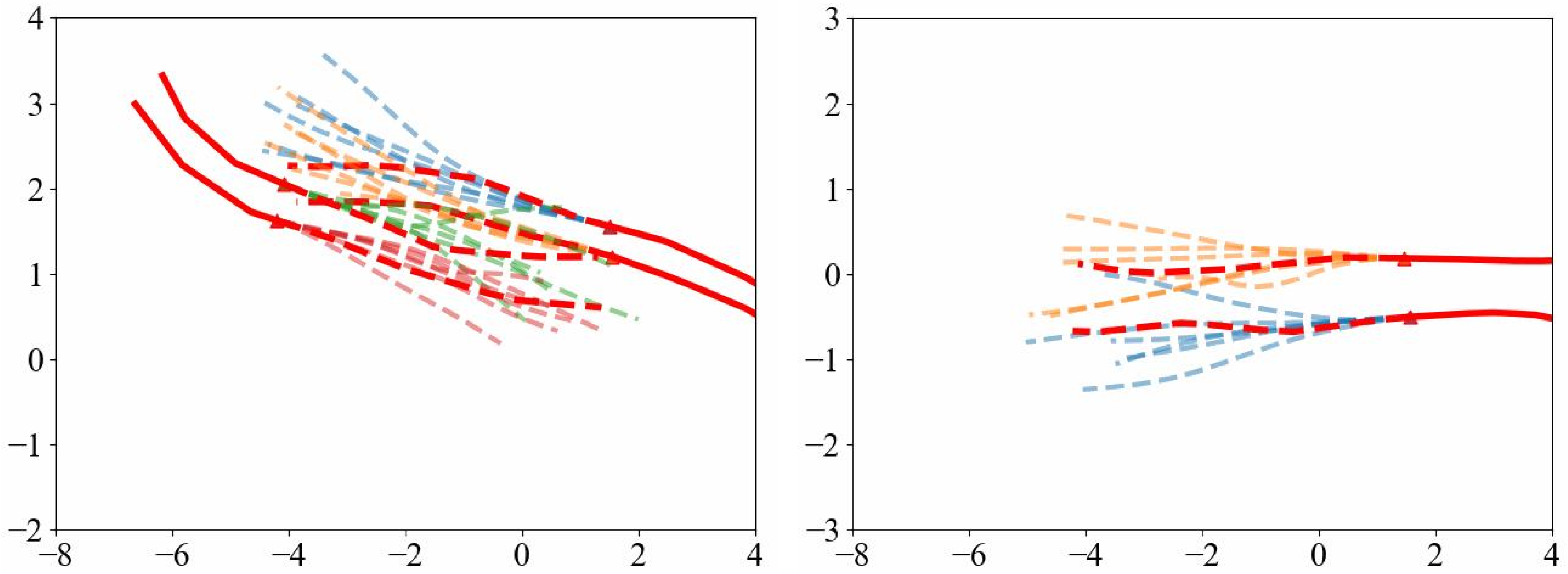
Input: $\mathcal{L}(\cdot,\cdot)$, $\phi(\cdot)$, c, $\{\bar{\alpha}_t, \beta_t, \lambda_t\}_{t=0}^{T-1}$,

$x_T \sim \mathcal{N}(0, \Sigma)$
for $t = T-1, \ldots, 1$ **do**
 $\varepsilon \sim \mathcal{N}(0, \Sigma)$
 $x_{t-1} = \frac{1}{\sqrt{\alpha_t}}(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}} \varepsilon_\theta(x_t, t)) + \sqrt{\beta_t}\varepsilon$
 $\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t - \sqrt{1-\bar{\alpha}_t}\varepsilon_\theta(x_t, t))$
 $g = -\nabla_{x_t} \mathcal{L}(\hat{x}_0, c)$
 $x_{t-1} = x_{t-1} + \lambda_t g$
 if RePaint **and** $t > 0$ **then**
  $\varepsilon' \sim \mathcal{N}(0, \Sigma)$
  $c_{t-1} = \sqrt{1-\beta_{t-1}} c_0 + \sqrt{\beta_{t-1}} \varepsilon$
  $\phi_{his}(x_{t-1}) = c_{t-1}$
 end if
end for
$y = \phi_{fut}(x_0)$

Output: y

Trajectory Prediction With Joint Metrics

- We consider using joint metrics that better reflects the significance of interaction modeling and scene-level accuracy



$$jointADE(Y,Y^*) = \frac{1}{TN}\min_{k=1}^{K}\sum_{n=1}^{N}\sum_{t=1}^{T}\left\|s_{t,n}-s^*_{t,n}\right\|$$

$$jointFDE(Y,Y^*) = \frac{1}{TN}\min_{k=1}^{K}\sum_{n=1}^{N}\left\|s_{T,n}-s^*_{T,n}\right\|$$

Method	minJADE / JFDE (m), K=20 ↓					
	ETH(1.4)	HOTEL(2.7)	UNIV(25.7)	ZARA1(3.3)	ZARA2(5.9)	ETH/UCY Avg.
S-GAN [28]	0.919 / 1.742	0.480 / 0.950	0.744 / 1.573	0.438 / 1.001	0.362 / 0.794	0.589 / 1.212
PECNet [29]	0.618 / 1.097	0.291 / 0.587	0.666 / 1.417	0.408 / 0.896	0.372 / 0.840	0.471 / 0.967
MemoNet[30]	0.499 / 0.859	0.222 / 0.416	0.686 / 1.466	0.349 / 0.723	0.385 / 0.864	0.428 / 0.866
Joint View Vertically [6]	0.652 / 0.839	0.186 / 0.309	0.523 / 1.091	0.331 / 0.634	0.267 / 0.547	0.392 / 0.684
Joint AgentFormer† [6]	0.543 / 0.883	0.211 / 0.377	0.596 / 1.247	0.309 / 0.612	0.282 / 0.584	0.388 / 0.741
Ours (GFTD)	0.505 / 0.873	0.174 / 0.297	0.649 / 1.305	0.340 / 0.667	0.308 / 0.620	0.395 / 0.752
Ours (GFTD + RePaint)	0.514 / 0.906	0.191 / 0.329	0.607 / 1.248	0.327 / 0.646	0.288 / 0.585	0.385 / 0.743

Pertubation-Robust Prediction

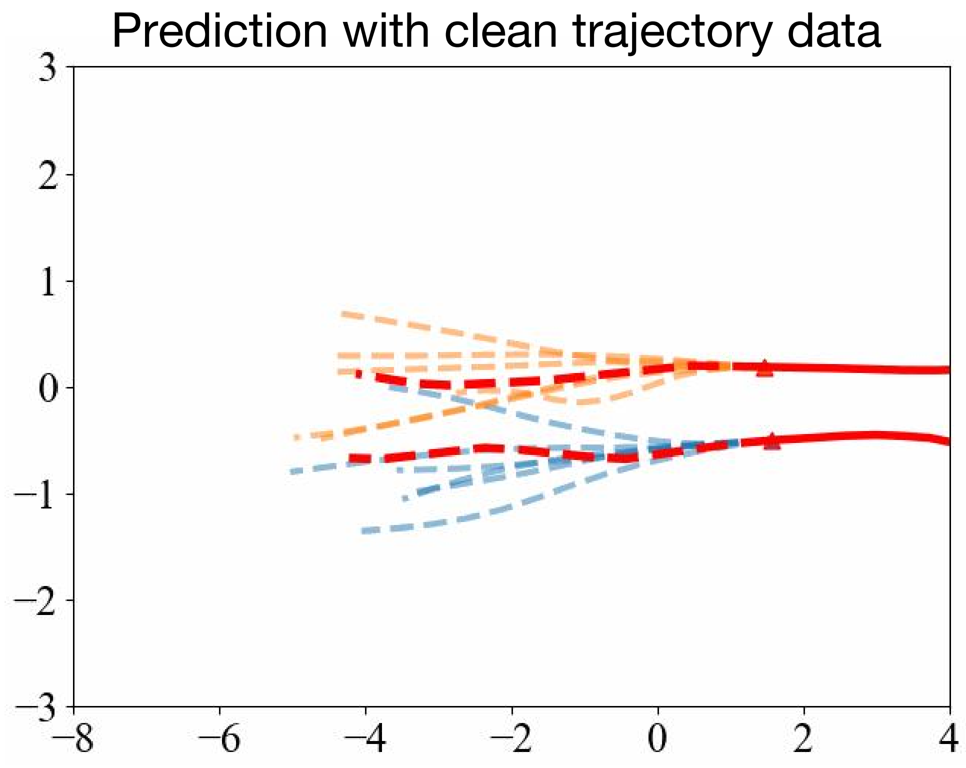
- Table III demonstrates noisy input prediction results, showcases that our model performs better at heavy noise levels.
- Table IV demonstrates model robustness against incomplete data.

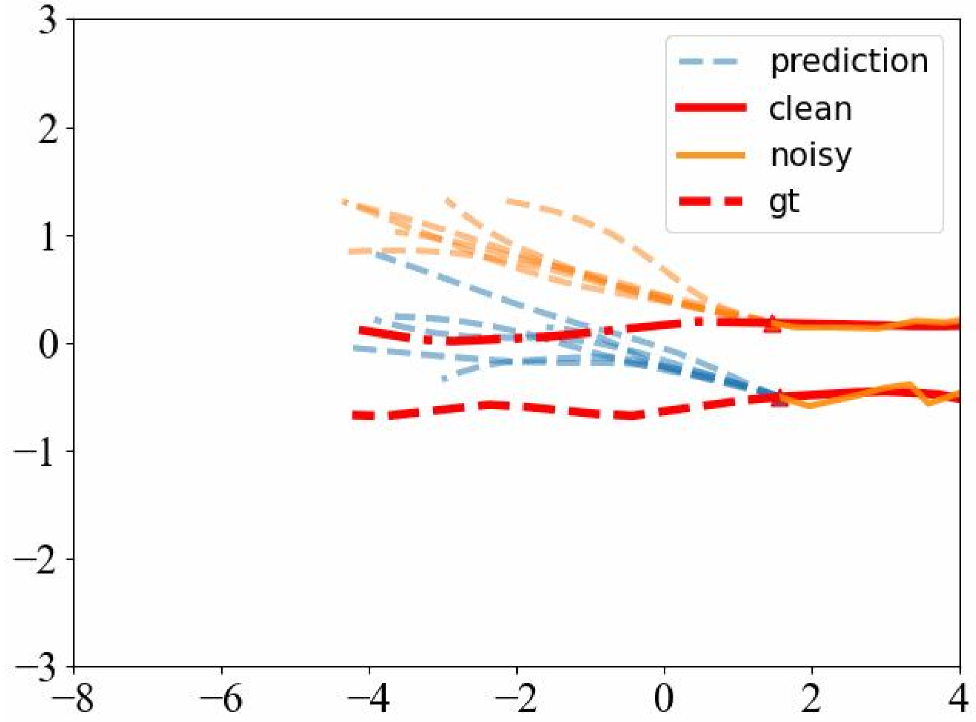
TABLE III: Prediction with Noisy Data

Noise Std	Model	minJADE / JFDE (m), K=20 ↓					
		ETH(1.4)	HOTEL(2.7)	UNIV(25.7)	ZARA1(3.3)	ZARA2(5.9)	ETH/UCY Avg.
0.00	Joint AgentFormer [5]	0.543 / 0.883	0.211 / 0.377	**0.596 / 1.247**	**0.309 / 0.612**	**0.282 / 0.584**	**0.388 / 0.741**
	Ours	**0.505 / 0.873**	**0.174 / 0.297**	0.649 / 1.305	0.340 / 0.667	0.308 / 0.620	0.395 / 0.752
0.05	Joint AgentFormer [5]	0.588 / **0.992**	0.253 / 0.433	**0.628 / 1.294**	**0.362 / 0.708**	0.361 / 0.714	**0.438** / 0.834
	Ours	**0.567** / 1.009	**0.198 / 0.334**	0.703 / 1.388	0.388 / 0.735	**0.341 / 0.675**	0.439 / **0.828**
0.15	Joint AgentFormer [5]	0.765 / 1.287	0.458 / 0.762	**0.834** / 1.642	0.649 / 1.254	0.736 / 1.312	0.688 / 1.254
	Ours	**0.636 / 1.076**	**0.283 / 0.463**	0.872 / **1.661**	**0.633 / 1.137**	**0.469 / 0.869**	**0.579 / 1.041**

TABLE IV: Prediction with Incomplete History Data

Missing ratio	Model	minJADE / JFDE (m), K=20 ↓					
		ETH(1.4)	HOTEL(2.7)	UNIV(25.7)	ZARA1(3.3)	ZARA2(5.9)	ETH/UCY Avg.
25%	Joint AgentFormer [5]	0.558 / 0.901	0.214 / 0.377	**0.624 / 1.286**	**0.336 / 0.653**	**0.303 / 0.614**	0.407 / 0.766
	Ours	**0.524 / 0.910**	**0.174 / 0.294**	0.662 / 1.321	0.341 / 0.670	0.314 / 0.627	**0.403 / 0.764**
50%	Joint AgentFormer [5]	0.619 / 1.022	0.219 / 0.378	**0.662 / 1.343**	0.374 / 0.721	0.327 / 0.649	0.440 / 0.823
	Ours	**0.511 / 0.897**	**0.176 / 0.302**	0.676 / 1.348	**0.347 / 0.677**	**0.321 / 0.642**	**0.406 / 0.773**
75%	Joint AgentFormer [5]	0.647 / 1.034	0.274 / 0.444	0.752 / 1.488	0.440 / 0.830	0.381 / 0.733	0.499 / 0.906
	Ours	**0.519 / 0.885**	**0.193 / 0.333**	**0.735 / 1.441**	**0.357 / 0.696**	**0.346 / 0.683**	**0.430 / 0.808**





Adding Gaussian noise to history trajectory

IROS '24
ABU DHABI

Controllable Generation

- Given desired goal point, GFTD can reliably generate desirable performance while maintancing both diversity and realism.

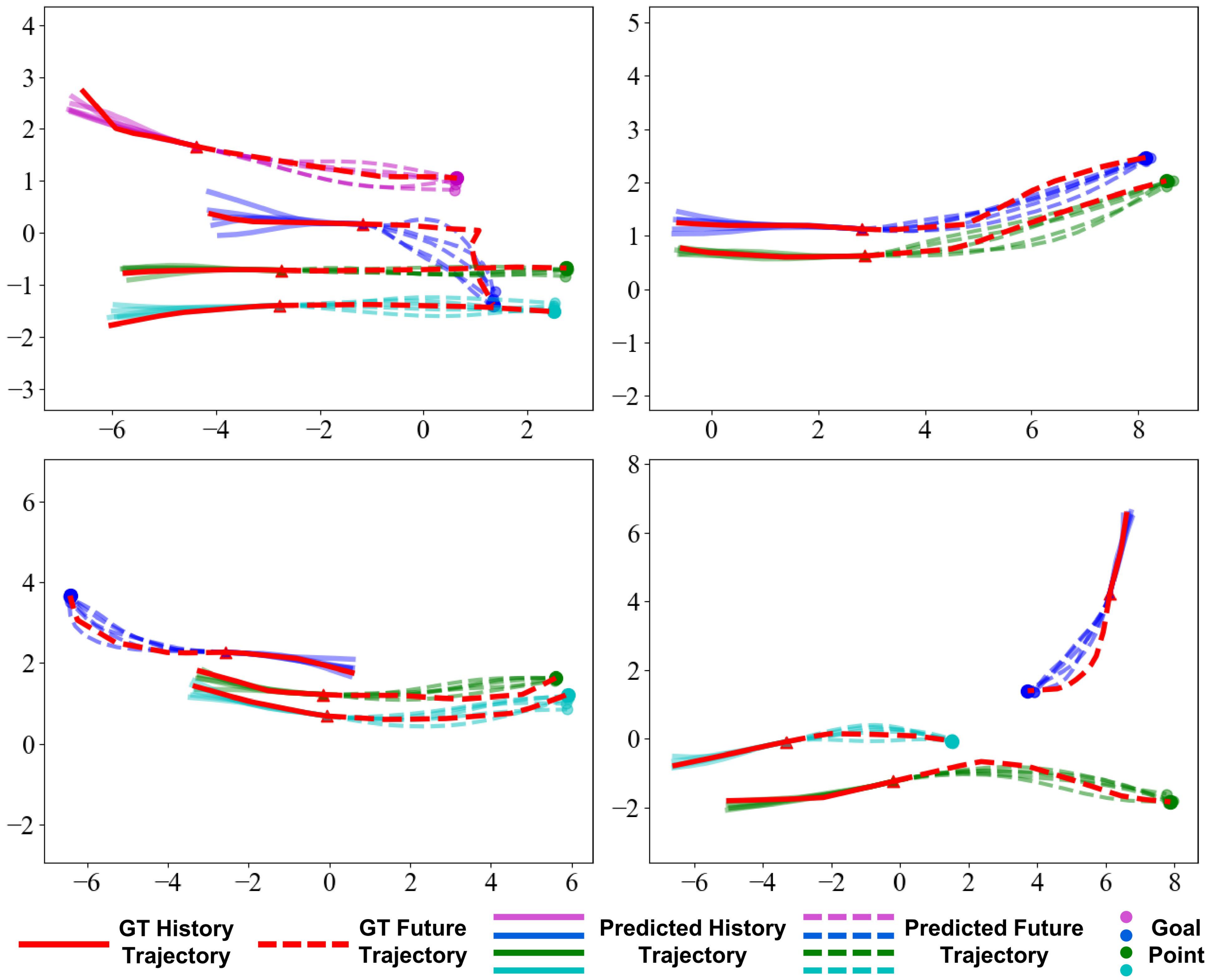


TABLE II: Controllable Generation Performance

Method	minJADE / JFDE (m), K=20 ↓					
	ETH(1.4)	HOTEL(2.7)	UNIV(25.7)	ZARA1(3.3)	ZARA2(5.9)	ETH/UCY Avg.
Baseline	0.505 / 0.873	0.174 / 0.297	0.649 / 1.305	0.340 / 0.667	0.308 / 0.620	0.395 / 0.752
Goal Point Guidance	0.224 / 0.064	0.069 / 0.033	0.280 / 0.394	0.103 / 0.035	0.094 / 0.032	0.154 / 0.112

IROS '24
ABU DHABI

Thank you!