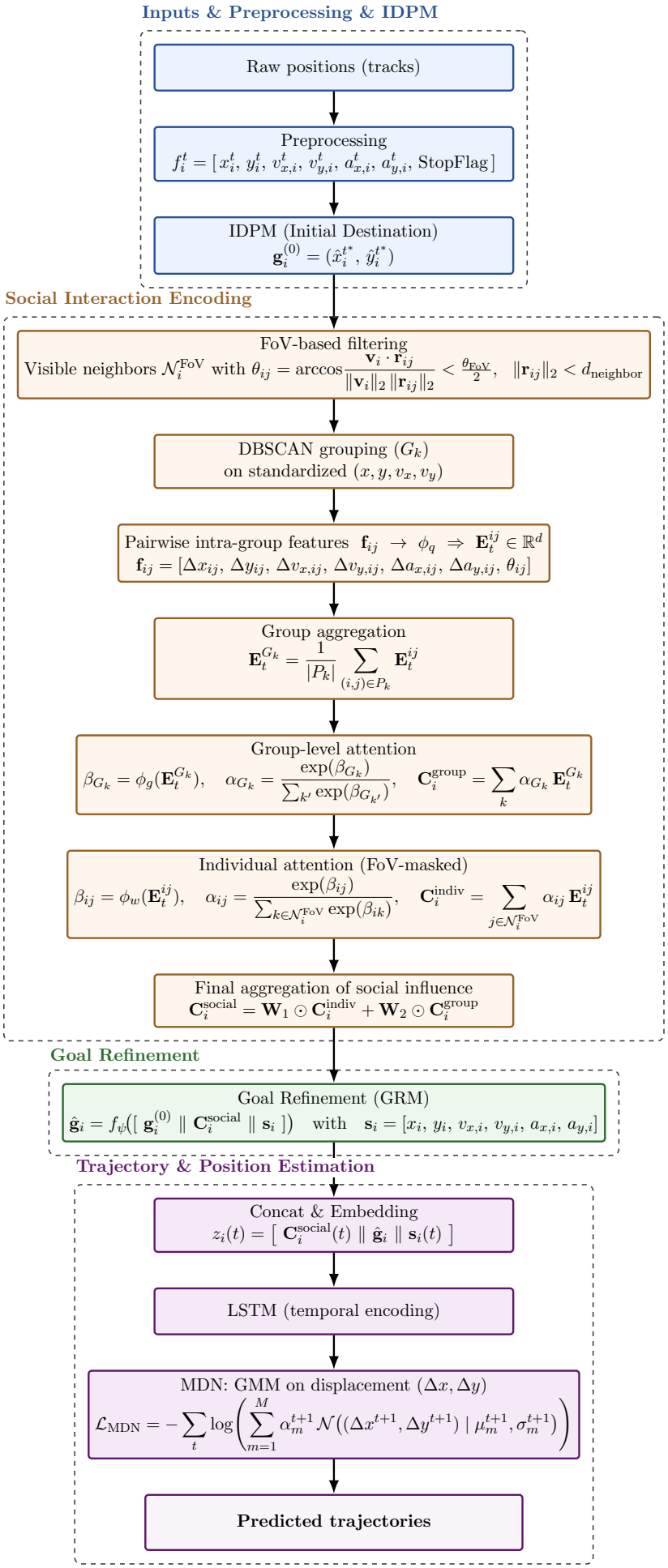
Inputs & Preprocessing & IDPM
Raw positions (tracks)
Preprocessing
$f_i^t = [\, x_i^t,\ y_i^t,\ v_{x,i}^t,\ v_{y,i}^t,\ a_{x,i}^t,\ a_{y,i}^t,\ \text{StopFlag}\,]$
IDPM (Initial Destination)
$\mathbf{g}_i^{(0)} = (\hat{x}_i^{t^*},\ \hat{y}_i^{t^*})$
Social Interaction Encoding
FoV-based filtering
Visible neighbors $\mathcal{N}_i^{\text{FoV}}$ with $\theta_{ij} = \arccos\dfrac{\mathbf{v}_i \cdot \mathbf{r}_{ij}}{\|\mathbf{v}_i\|_2\,\|\mathbf{r}_{ij}\|_2} < \dfrac{\theta_{\text{FoV}}}{2},\quad \|\mathbf{r}_{ij}\|_2 < d_{\text{neighbor}}$
DBSCAN grouping (G_k)
on standardized (x, y, v_x, v_y)
Pairwise intra-group features $\mathbf{f}_{ij} \rightarrow \phi_q \Rightarrow \mathbf{E}_t^{ij} \in \mathbb{R}^d$
$\mathbf{f}_{ij} = [\Delta x_{ij},\ \Delta y_{ij},\ \Delta v_{x,ij},\ \Delta v_{y,ij},\ \Delta a_{x,ij},\ \Delta a_{y,ij},\ \theta_{ij}]$
Group aggregation
$\mathbf{E}_t^{G_k} = \dfrac{1}{|P_k|} \sum_{(i,j)\in P_k} \mathbf{E}_t^{ij}$
Group-level attention
$\beta_{G_k} = \phi_g(\mathbf{E}_t^{G_k}),\quad \alpha_{G_k} = \dfrac{\exp(\beta_{G_k})}{\sum_{k'} \exp(\beta_{G_{k'}})},\quad \mathbf{C}_i^{\text{group}} = \sum_k \alpha_{G_k}\, \mathbf{E}_t^{G_k}$
Individual attention (FoV-masked)
$\beta_{ij} = \phi_w(\mathbf{E}_t^{ij}),\quad \alpha_{ij} = \dfrac{\exp(\beta_{ij})}{\sum_{k\in\mathcal{N}_i^{\text{FoV}}} \exp(\beta_{ik})},\quad \mathbf{C}_i^{\text{indiv}} = \sum_{j\in\mathcal{N}_i^{\text{FoV}}} \alpha_{ij}\, \mathbf{E}_t^{ij}$
Final aggregation of social influence
$\mathbf{C}_i^{\text{social}} = \mathbf{W}_1 \odot \mathbf{C}_i^{\text{indiv}} + \mathbf{W}_2 \odot \mathbf{C}_i^{\text{group}}$
Goal Refinement
Goal Refinement (GRM)
$\hat{\mathbf{g}}_i = f_\psi([\ \mathbf{g}_i^{(0)} \parallel \mathbf{C}_i^{\text{social}} \parallel \mathbf{s}_i\ ])\quad \text{with}\quad \mathbf{s}_i = [x_i,\ y_i,\ v_{x,i},\ v_{y,i},\ a_{x,i},\ a_{y,i}]$
Trajectory & Position Estimation
Concat & Embedding
$z_i(t) = [\ \mathbf{C}_i^{\text{social}}(t) \parallel \hat{\mathbf{g}}_i \parallel \mathbf{s}_i(t)\ ]$
LSTM (temporal encoding)
MDN: GMM on displacement $(\Delta x, \Delta y)$
$\mathcal{L}_{\text{MDN}} = -\sum_t \log\left(\sum_{m=1}^{M} \alpha_m^{t+1}\, \mathcal{N}\big((\Delta x^{t+1}, \Delta y^{t+1}) \mid \mu_m^{t+1}, \sigma_m^{t+1}\big)\right)$
Predicted trajectories