Appendix

A More experimental details

To sample a heterogeneous few-shot task for a specific action, we first select an induced subgraph of the original human skeleton, as described in Section 4.1. Each task consists of five labeled support instances for training and two query instances for testing. During meta-testing, all instances are randomly sampled from actor 5. During meta-training, the instances are sampled randomly from the remaining actors, as in previous works [8, 27, 26]. We train the models for 500 epochs, with 50 meta-batches sampled per epoch. Each meta-batch consists of one task per meta-training action, resulting in a total of 275,000 tasks seen during meta-training. For evaluation, we randomly sample 500 tasks per meta-test action from actor 5.

B Cross-validation

We conducted 5-fold cross-validation across actions to assess our model's performance for all actions shown in Table 5. The splits were designed to be intuitively more difficult by grouping similar actions in the same fold, such as walking, walking dog, and walking together. As expected, the results for walking degraded compared to the previous experimental setup proposed by [17].

	GrapHetNet				
Actions (by fold)	80	160	320	400	Avg
walkingdog	0.409	0.747	1.186	1.38	0.93
walkingtogether	0.251	0.50	0.74	0.82	0.58
walking	0.17	0.32	0.50	0.69	0.42
smoking	0.233	0.23	0.51	0.64	0.40
eating	0.209	0.28	0.60	0.71	0.45
posing	0.212	0.28	0.49	0.67	0.41
phoning	1.26	2.53	2.83	2.98	2.40
directions	0.61	1.12	1.43	1.5	1.17
waiting	0.257	0.39	0.77	0.90	0.58
discussion	0.19	0.46	0.97	1.25	0.72
takingphoto	0.21	0.57	0.98	1.20	0.74
sittingdown	0.24	0.66	1.02	1.19	0.78
greeting	0.343	0.71	1.33	1.71	1.02
purchases	0.35	0.49	0.85	0.93	0.65
sitting	0.32	0.40	0.67	0.67	0.52

Table 5. Results for 5-fold cross-validation for GraphHetNet. Folds are separated by horizontal line.

C Forecast with full past horizon

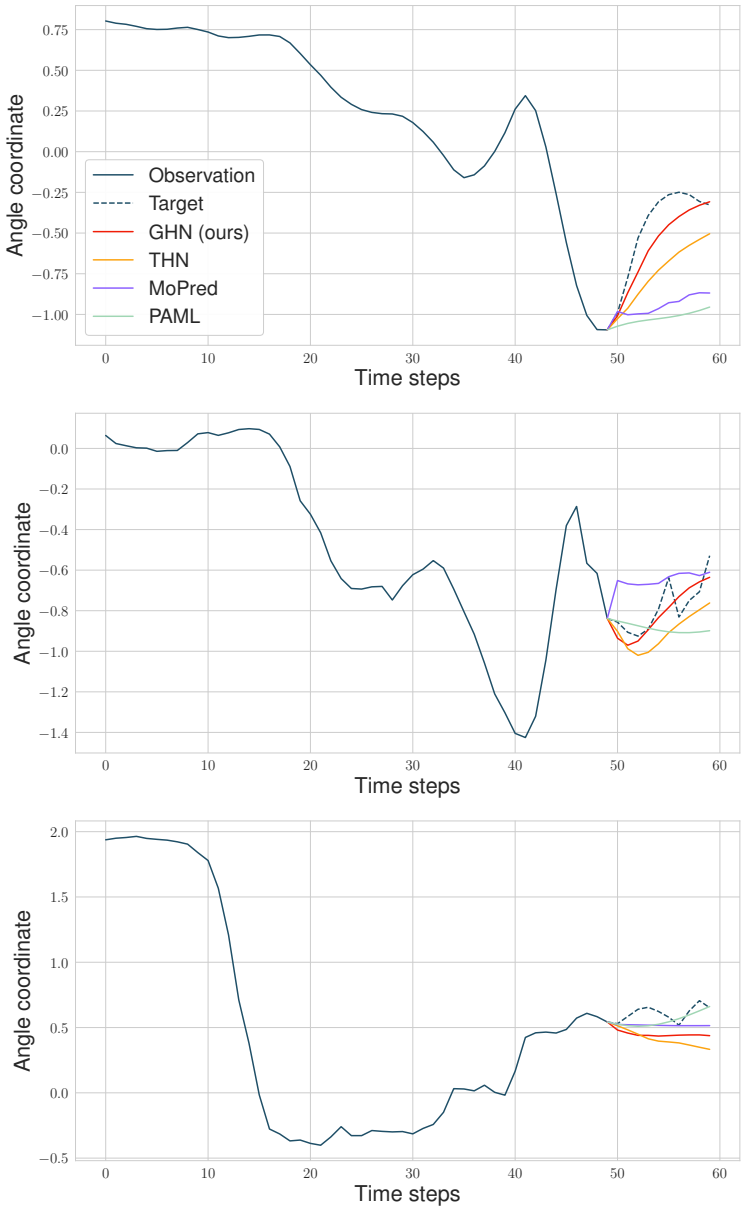

Fig. 5. A full zoomed out version of Figure 1: Examples of three motion predictions in the exponential map for GHNand baseline approaches. We sampled two examples where our approach (red) has the lowest error and one where a baseline performs best.