End-To-End Semi-supervised Learning for Differentiable Particle Filters

Hao Wen, Xiongjie Chen, Georgios Papagiannis, Conghui Hu, and Yunpeng Li

Abstract—Recent advances in incorporating neural networks into particle filters provide the desired flexibility to apply particle filters in large-scale real-world applications. The dynamic and measurement models in this framework are learnable through the differentiable implementation of particle filters. Past efforts in optimising such models often require the knowledge of true states which can be expensive to obtain or even unavailable in practice. In this paper, in order to reduce the demand for annotated data, we present an end-to-end learning objective based upon the maximisation of a pseudo-likelihood function which can improve the estimation of states when large portion of true states are unknown. We assess performance of the proposed method in state estimation tasks in robotics with simulated and real-world datasets.

I. INTRODUCTION

Sequential state estimation task, which involves estimating unknown state from a sequence of observations, finds a variety of applications including target tracking [1], [2], navigation [3], [4], and signal processing [5], [6]. Recursive Bayesian filtering is a probabilistic approach to sequential state estimation [7], which requires the specification of a dynamic model that describes how the hidden state evolves over time and a measurement model that defines the likelihood of the observation given the estimated state.

Algorithms for solving such a problem involve constructing these probabilistic models or learning them from data [8], [9]. Estimation of the models’ parameters is often known as parameter inference, which normally comprises either Bayesian or Maximum Likelihood (ML) methods [10], [11]. These methods can also be classified as offline or online depending on whether the data is processed in batch [12], [13] or recursively [11], [14]. Online approach is often preferable with streaming data, especially if the amount of data to be processed is large [14]. While online Bayesian methods suffer from degeneracy problems, maximum likelihood-based methods, including direct optimisation approaches, gradient ascent, Expectation Maximisation (EM), have shown promises in parameter inference with large datasets [11].

However, the aforementioned methods assume that the structures or part of parameters of the dynamic and measurement models are known [11]. Yet, formulating the dynamic model and measurement model is challenging, especially when the hidden state and observation data are in high-dimensional space or there is a lack of prior knowledge about the dynamic and measurement models. For example, sophisticated models are often required to describe the characteristics exhibit in complex, heterogeneous datasets such as streaming videos, text and medical records [15]. For these tasks, it is often impractical to specify dynamic and measurement models a priori [16].

Recent advances in incorporating neural networks into Bayesian filtering provide the desired flexibility to learn the dynamic and measurement models and their parameters. Traditional Bayesian filtering methods can be implemented as differentiable filters, i.e. the dynamic and measurement models can be specified through neural networks and their parameters can be learned through backpropagation. Differentiable filters have been implemented for Kalman filters [17], [18], [16], histogram filters [19], and particle filters [20], [21], [22], [23]. The incorporation of neural networks to particle filters, a.k.a. sequential Monte Carlo methods, is especially appealing due to the nonlinear nature of neural networks and the ability of particle filters to track posterior distribution with non-Gaussian models. However, to the best of our knowledge, optimisation of all of these differentiable filters requires sufficient training data with true state information or labels during the training stage, which can be expensive to obtain or even inaccessible in real-world scenarios.

To this end, we present the semi-supervised differentiable particle filter (SDPF) which leverages the unlabelled data in addition to a small subset of labelled data samples. This semi-supervised learning setting is of immense practical interest for differentiable particle filters in a wide range of applications where the unlabelled data are abundant and the access to labels is limited [24], [25], [26]. Our main contributions are three-folds: (i) We introduce an end-to-end learning objective based upon the maximisation of a pseudo-likelihood function for semi-supervised parameter inference in the differentiable particle filtering framework; (ii) We specify general dynamic and measurement model structures that can facilitate the optimisation of differentiable particle filters while maintaining the flexibility of models; (iii) We demonstrate superior performance of the proposed semi-supervised differentiable particle filters through experiments with simulated and real-world datasets.

The rest of this paper is organised as follows. We present the problem statement in Section II. A discussion of the background and related work is provided in Section III. We introduce the proposed method in Section IV. Section V provides details of experiments to evaluate the performance of proposed method. We conclude the paper in Section VI.

The authors are with Department of Computer Science, University of Surrey, Surrey GU2 7XH, United Kingdom, e-mail: {h.wen, xiongjie.chen, g.papagiannis, conghui.hu, yunpeng.li}@surrey.ac.uk.
II. PROBLEM STATEMENT

Recursive Bayesian filtering involves tracking the marginal posterior distribution \( p(s_t|o_{1:t}, a_{1:t}) \), where \( s_t \) is the hidden state at current time \( t \), \( o_{1:t} = \{o_1, o_2, \ldots, o_t\} \) and \( a_{1:t} = \{a_1, a_2, \ldots, a_t\} \) are the history of observations and actions up to time step \( t \), respectively. Bayesian filtering is considered in general probabilistic state space models where the dynamic and measurement models can be described as:

\[
\begin{align*}
    s_1 &\sim \mu(s_1) , \\
    s_t &\sim g_0(s_t|s_{t-1}, a_t) \quad \text{for} \ t \geq 2 , \\
    o_t &\sim l_\theta(o_t|s_t) \quad \text{for} \ t \geq 2 .
\end{align*}
\]

\( \mu(s_1) \) is the stationary distribution of the hidden state at initial time, \( g_0(s_t|s_{t-1}, a_t) \) is the dynamic model which describes the transition of hidden state, \( l_\theta(o_t|s_t) \) is the measurement model which describes the relation between the observation \( o_t \) and the hidden state \( s_t \). We use \( \theta \) to denote all parameters of the dynamic and measurement models. In this paper, we focus on estimating the parameters \( \theta \) of the dynamic and measurement models parameterised by neural networks.

III. BACKGROUND AND RELATED WORK

In this section, we briefly review how the bootstrap particle filter and differentiable particle filters are used for recursive Bayesian filtering.

A. Bootstrap particle filter

Particle filters estimate the posterior distribution of hidden state by a set of weighted samples \( \{(s^i_t, w^i_t)\}_{i=1}^{N_p} \), where \( s^i_t \) is the state value of the \( i \)-th particle at time step \( t \), \( w^i_t \) is the corresponding particle weight and \( \sum_{i=1}^{N_p} w^i_t = 1 \), and \( N_p \) represents the number of particles. In the bootstrap particle filter [27], the dynamic model and measurement model are used to propagate the particles and update the associated weights, respectively, which can be described by the following models:

\[
\begin{align*}
    s^1 &\sim \mu(s_1) , \\
    s^i_t &\sim g_0(s^i_t|s^i_{t-1}, a_t) \quad \text{for} \ t \geq 2 , \\
    l^i_t &\sim l_\theta(o_t|s^i_t) \quad \text{for} \ t \geq 2 ,
\end{align*}
\]

where \( g_0(\cdot|\cdot, \cdot) \) is the dynamic model which describes the transition of the hidden state, \( s^i_t \) is the predicted particle state for the \( i \)-th particle at time step \( t \), and \( l^i_t \) is the observation likelihood. \( \theta \) denotes all parameters of the dynamic and measurement models. The particle weights are updated recursively according to the observation likelihood:

\[
    w^i_t = \zeta w^i_{t-1} l^i_t ,
\]

where \( \zeta \) is the normalization factor and \( \zeta^{-1} = \sum_{i=1}^{N_p} w^i_{t-1} l^i_t \). Particles may be resampled proportional to their weights to avoid particle degeneracy.

B. Differentiable particle filters

In most variants of particle filters, the dynamic model \( g_0(s_t|s_{t-1}, a_t) \) and measurement model \( l_\theta(o_t|s_t) \) are assumed to be known. In order to remove this restriction, differentiable particle filters [20, 21] are proposed to learn the parameters \( \theta \) of models \( g_0(s_t|s_{t-1}, a_t) \) and \( l_\theta(o_t|s_t) \) from data. Differentiable particle filters incorporate neural network into Bayesian inference where the learnable dynamic and measurement models can be optimised through gradient descent. As the sampling and resampling operations are not differentiable, the reparameterization trick was proposed in [28] to perform the sampling operation in a differentiable fashion. Specifically, the reparameterization trick constructs the prior distribution of the hidden state by employing a deterministic and differentiable function w.r.t its input and a noise vector.

The architecture design of dynamic and measurement models for differentiable particle filters varies in different applications. For the design of the dynamic model, in [20] and [21], the relative motion of hidden state is learned by transforming actions in the local coordinate to the global coordinate through a neural network. For measurement model, [20] adopts a neural network with observation features as input which directly outputs the likelihood of predicted particles. In [21], the measurement model is learned by incorporating the features of observation and a global map. In [22], both dynamic and measurement models are learned using the recurrent neural network (RNN) architecture.

In order to achieve differentiable resampling, the backpropagation of gradient is truncated during resampling at every time step [20]. Besides, [21] introduced soft resampling such that the gradient can be backpropagated from the weight of the sampled ancestors. In [29], a learned neural network resampler, which transforms the particle sets by a permutation-invariant encoder-decoder network, was introduced and resampled particles are aggregated by pooling.

The optimisation of existing differentiable particle filters [20, 21, 23] relies on the knowledge of true states for all time steps. In [20], the parameters of the differentiable particle filters are learned by minimising the negative log-likelihood (NLL) of the true state under the estimated posterior distribution. In [21] and [22], differentiable particle filters are trained by minimising the mean squared error (MSE) between the predicted state and the ground truth state.

IV. SEMI-SUPERVISED LEARNING FOR DIFFERENTIABLE PARTICLE FILTERS

To alleviate the requirements for annotated data, we formulate a semi-supervised learning framework for training differentiable particle filters and introduce specifically designed dynamic and measurement models to facilitate the learning procedure.

Dynamic model. For efficient and stable gradients calculation between time steps, we adopt the practice in [20] and [21] where the relative motion of hidden state between time steps is modelled. The dynamic model is formulated as:

\[
    s^i_t = g_0(s^i_t|s^i_{t-1}, a_t) = s^i_{t-1} + f_\theta(s^i_{t-1}, a_t) + \epsilon^i .
\]
where $\theta$ denotes the parameters of a neural network $f_\theta$ which transforms actions in the local coordinate to the global coordinate and models the relative motion between time steps, and $c^i$ is an auxiliary noise vector with independent marginal $p(\epsilon)$ used in the reparameterization trick [23].

**Measurement model.** The measurement model in Eq. (4) defines the likelihood of the observation given the predicted state. The measurement model can be regarded as a generative model which provides the distribution of observation data. However, if the observation data are in high-dimensional spaces, it is challenging to directly formulate the measurement model effectively. In such cases, an encoder network can be used to learn a more compact representation of observation data for dimension reduction. We approximate the likelihood used to learn a more compact representation of observation spaces, it is challenging to directly formulate the measurement model.

The measurement model in Eq. (6) can be regarded as a generative network $\hat{h}_\theta : s_i \mapsto e_t$, which extracts features $\hat{e}_t$ from predicted particle states $\hat{s}_i$, is employed as the state encoder. In summary, the measurement model $l_\theta(\theta,\theta)$ can be expressed as follows:

$$e_t = h_\theta(o_t) ,$$

$$\hat{e}_t = \hat{h}_\theta(s_i) ,$$

$$l_t = l_\theta(o_t,\hat{s}_i) = d_\theta(\hat{e}_t,e_t) ,$$

where $d_\theta(\hat{e}_t,e_t)$ measures the similarity between the features $\hat{e}_t$ and $e_t$, and higher likelihood $l_t$ implies smaller distance between $\hat{e}_t$ and $e_t$. Specifically, $d_\theta(\hat{e}_t,e_t)$ is defined as:

$$d_\theta(\hat{e}_t,e_t) = 1/c(\hat{e}_t,e_t) ,$$

$$c(\hat{e}_t,e_t) = 1 - \frac{\hat{e}_t \cdot e_t}{\|\hat{e}_t\|_2 \|e_t\|_2} ,$$

where $c(\hat{e}_t,e_t)$ calculates the cosine distance between $\hat{e}_t$ and $e_t$.

**Semi-supervised learning.** We derive a learning objective for semi-supervised learning based on maximising a pseudo-likelihood. We start by defining a logarithmic pseudo-likelihood $Q_{semi}(\theta)$:

$$Q_{semi}(\theta) = \sum_{b=0}^{m-1} \log p_\theta(O_b|A_b) ,$$

where $O_b = o_{bL+1:(b+1)L}$ and $A_b = a_{bL+1:(b+1)L}$ are the $b$-th block of observations and the $b$-th block of actions respectively. $m$ denotes the number of blocks and $L$ is the block length. The likelihood of the $b$-th block of observations $p_\theta(O_b|A_b)$ can be expressed by:

$$p_\theta(O_b|A_b) = \int_{s_L} p_\theta(S_b, O_b|A_b)dS_b$$

$$= \int_{s_L} p_\theta(S_b|A_b)p_\theta(O_b|S_b, A_b)dS_b ,$$

where $S_b = s_{bL+1:(b+1)L}$ is the $b$-th block of states.

Due to the Markovian property of the dynamic model as in Eq. (5) and the measurement model described via Eq. (6), $p_\theta(S_b|A_b)$ and $p_\theta(O_b|S_b, A_b)$ can be expressed as:

$$p_\theta(S_b|A_b) = \mu_\theta(s_{bL+1}) \prod_{m=bL+2}^{(b+1)L} g_\theta(s_m|s_{m-1}, a_m) ,$$

$$p_\theta(O_b|S_b, A_b) = \prod_{m=bL+1}^{(b+1)L} l_\theta(a_m|s_m) .$$

According to the ergodicity assumption in employing pseudo-likelihood [14], the average of logarithmic pseudo-likelihood satisfies:

$$\lim_{m \to \infty} \frac{1}{m} \sum_{b=0}^{m-1} \log p_\theta(O_b|A_b) = Q_{semi}(\theta) ,$$

where $Q_{semi}(\theta)$ is defined as:

$$Q_{semi}(\theta) = \int_{s_L \times o_L} \log p_\theta(O|A) \cdot p_\theta(A)dSdO ,$$

and $S$ is a block of states, and $\theta$ is the parameter value estimated at the $b$-th block.

With the differentiable particle filtering framework, we can approximate $Q(\theta, \theta)$ with $Q(\theta, \theta)$ as follows to form the optimisation objective for samples without true labels:

$$Q(\theta, \theta) = \sum_{i=1}^{N_p} w_b^i \log p_\theta(S_b^i, O_b|A_b)$$

$$= \sum_{i=1}^{N_p} w_b^i \log (\mu_\theta(s_{bL+1})l_\theta(o_{bL+1}|s_{bL+1})$$

$$\prod_{m=bL+1}^{(b+1)L} g_\theta(s_m|s_{m-1}, a_m)l_\theta(a_m|s_m) ,$$

where $w_b^i$ is the particle weight of the $i$-th particle at block $b$. $\mu_\theta(s_{bL+1})$ is the stationary distribution of the particle state at step $bL + 1$.

Based on the logarithmic pseudo-likelihood function above, we now propose the learning objective for semi-supervised learning as follows:

$$\theta = \arg \min_{\theta \in \Theta} \lambda_1 L(\theta) - \lambda_2 Q(\theta) ,$$

$$L(\theta) = \frac{1}{|T_1|} \sum_{t \in T_1} \delta(s_t^*, s_t^*) ,$$

$$Q(\theta) = \frac{1}{m} \sum_{b=0}^{m-1} Q(\theta, \theta) ,$$
where $\lambda_1, \lambda_2$ are the scaling factors depending on the
magnitude of components $L(\theta)$ and $Q(\theta)$. $T_1$ is the set of time
with true labelled states and $|T_1|$ is the number of elements in
$T_1$. $\delta(s^*_t, \hat{s}_i)$ is a loss function between true labelled states $s^*_t$
and estimated states $\hat{s}_i = \sum_{l=1}^{N_p} w^l_t s^l_t$. $L(\theta)$ is the objective
for supervised learning, e.g. the mean squared error (MSE) [21]:

$$L(\theta) = \frac{1}{|T_1|} \sum_{t \in T_1} \delta(s^*_t, \hat{s}_i) = \frac{1}{|T_1|} \sum_{t \in T_1} (s^*_t - \hat{s}_i)(s^*_t - \hat{s}_i)^T,$$

(26)

With the derived semi-supervised learning objective, the pseudo-
code for the semi-supervised differentiable particle filters (SDPF)
framework is presented in Algorithm 1.

V. EXPERIMENTS

In this section, we investigate the performance of semi-
supervised differentiable particle filters (SDPF) where ground
truth information is unavailable for a large portion of time
steps. The same experiment environments as in [20] and
[21] with simulated and real-world datasets, namely simulated
Maze environment [30] and House3D environment [31], are
used to evaluate algorithms. Since the global map consisting
in the measurement model in [21] is unavailable in the Maze
environment, to make a fair comparison focusing on the semi-
supervised setting, the proposed method is compared with
[20] and [21] in the Maze environment and the House3D
environment, respectively. In both environments, the perfor-
mance of long short-term memory (LSTM) network [32]
is also compared.

A. Simulated maze environment

1) Experiment setup: In the first experiment, the SDPF is
tested on the task of robot global localisation in a simulated
maze environment. It is modified from the DeepMind Lab
environment [30] where unique wall textures are removed to
make the hidden state to be partially observable. There are
three different sizes of mazes in this environment, namely
Maze 1, Maze 2, and Maze 3 with the size ranging from
small to large. The goal of this task is to estimate the state
of the robot from a sequence of visual images and odometry
measurements. The robot randomly walks through the maze,
and the observation images come from first-person viewpoint
of the robot with $32 \times 32$ RGB image size and the actions
data come from the odometry measurements.

Both the training and testing data consist of 1000 trajectories
of 100 steps with one step per second. In the semi-supervised
setting, we take 100 trajectories for training with 10% of
randomly chosen time steps provided with ground truth state
values, and validate on the remaining 900 trajectories. The
states $s = [s_x, s_y, s_\beta]$ and actions $a = [a_x, a_y, a_\beta]$ represent the
coordinate of position $(x, y)$ and orientation angle $\beta$, and
the velocity of translation and rotation, respectively.

The particles are initialised uniformly across the maze.
The number of particle used in the SDPF and DPF is set to be
100. In the dynamic model Eq. (3), the actions $a_t$ come from
odometry measurements which describe the relative motion

| Algorithm 1: Semi-supervised differentiable particle filters framework |
|---|
| **Input:** |
| $o_{1:T}$ | Observations | $\mu_0(s_1)$ | Initial distribution of $s_1$ |
| $a_{1:T}$ | Actions | $N_p$ | Particle number $T$ |
| $\alpha$ | Learning rate | $\hat{s}_i$ | True labelled state |
| $\lambda_1, \lambda_2$ | Scaling factors | $N_{thres}$ | Resampling threshold |
| $g_0$ | Dynamic model | $L$ | Block length |
| $l_\theta$ | Meas. model | $f_\theta$ | Action transformer |
| $d_\theta$ | Cosine distance | $\delta(\cdot)$ | Supervised loss function |
| $h_\theta$ | Obs. encoder | $\epsilon$ | Reparam. noise vector |
| $\tilde{h}_\theta$ | State encoder | $p(\epsilon)$ | Distribution of $\epsilon$ |

1 Initialise parameters $\theta$ of $f_\theta$, $h_\theta$ and $\tilde{h}_\theta$ randomly;
2 Draw particles $\{s^*_t\}_{i=1}^{N_p}$ from $\mu_0(s_1)$;
3 Set particle weights $\{w^l_1\}_{i=1}^{N_p} = \frac{1}{N_p}$;
4 Set $Q = 0$ and $b = 0$;
5 while $\theta$ not converged do
   for $k = 2$ to $T$ do
      Compute the effective sample size:
      $$N_{eff} = \frac{1}{\sum_{i=1}^{N_p} (w^l_{k-1})^2};$$
      if $N_{eff} < N_{thres}$ then
         Select ancestor index $A^i_{k-1}$ with
         $$Pr(A^i_{k-1} = i) = w^l_{k-1} \text{ for } i = 1, ..., N_p;$$
         $$\{w^l_{k-1}\}_{i=1}^{N_p} = \frac{1}{N_p};$$
      else
         $$A^i_{k-1} = i \text{ for } i = 1, ..., N_p;$$
      end
      $e_k = h_\theta(a_k);$$
      for $i = 1, ..., N_p$ do
         Draw particles $s^i_k \sim g_0(s_k|s^i_{k-1}, a_k) = s^i_{k-1} + f_\theta(s^i_{k-1}, a_k) + \epsilon$ where $\epsilon \sim p(\epsilon);$ | $|n_i = \mu_0(s^i_{L+1})h_\theta(s^i_{m+1}|s^i_{m-1}, a_m)|f_\theta(o_m|s^i_m);$ |
         $\hat{e}_k^i = h_\theta(s^i_k), l^i_k = d_\theta(\hat{e}_k^i, e_k);$ | $|n_i = \mu_0(s^i_{L+1})h_\theta(s^i_{m+1}|s^i_{m-1}, a_m)|f_\theta(o_m|s^i_m);$ |
         $w^i_k = l^i_k w^i_{k-1};$ | $|n_i = \mu_0(s_{L+1})h_\theta(s_{m+1}|s_{m-1}, a_m)|f_\theta(o_m|s^i_m);$ |
         Set $s^i_{1:k} = (s^i_{1:k-1}, s^i_k);$ | $|n_i = \mu_0(s_{L+1})h_\theta(s_{m+1}|s_{m-1}, a_m)|f_\theta(o_m|s^i_m);$ |
      end
      Set $s^i_{1:k}$ for $i = 1, ..., N_p$; |
      Normalize weights $\{w^i_k\}_{i=1}^{N_p}$ so that
      $$\sum_{i=1}^{N_p} w^i_k = 1;$$ |
      if $k \mod L = 0$ then
         for $i = 1, ..., N_p$ do
            Compute $Q = Q + \sum_{i=1}^{N_p} w^i_k \log(n_i);$
            $b = b + 1;$ |
         end
         Total loss:
         $$\mathcal{L} = \lambda_1 \frac{1}{|T_1|} \sum_{t \in T_1} \delta(s^*_t, \sum_{i=1}^{N_p} w^l_t s^l_t) - \lambda_2 \frac{1}{Q} Q;$$ |
         Update model parameters $\theta \leftarrow \theta - \alpha \nabla_\theta \mathcal{L};$ |
      end
   end
end

1 Code is available at https://github.com/HaoWen-Surrey/SemiDPF.
Fig. 1. The Root Mean Square Error (RMSE) comparison using SDPF, DPF and LSTM in Maze 1, Maze 2 and Maze 3 (top, middle, bottom row, respectively). (a): The mean and standard deviation of RMSE at the last time step on validation data during training starting from epoch 1. (b): The boxplot of the RMSE at the specified time steps evaluated at 1000 trajectories on testing data. The percentage of labelled data corresponding to results presented in (a), (b) and (c) is 10%. (c): The RMSE for 1000 trajectories on testing data with different percentage of available ground truth data. The error bar represents the standard error of the reported RMSE.

in robot coordinate frame, and are transformed into global coordinate frame \( f_0(s_{t-1}, a_t) \). Here, the predicted particles are sampled from the differentiable function \( \tilde{s}_t = s_{t-1} + f_0(s_{t-1}, a_t) + \text{diag}(\sigma_x, \sigma_y, \sigma_\beta) \zeta_t \), where \( \zeta_t \sim \mathcal{N}(0; I) \) is a noise vector from a standard multivariate Gaussian and \( I \) is an identity matrix. The standard deviations \( \sigma_x, \sigma_y, \) and \( \sigma_\beta \) for translation in \( x \) and \( y \) axes and the rotation are set to be 20.0, 20.0, and 0.5, respectively. For SDPF, we choose the block length \( L = 20 \) time steps. Adam \cite{Kingma2014AdamAM} is used as the optimiser and the learning rate is set to 0.0003.

2) Experimental results: In this experiment, the error metric is the Root Mean Square Error (RMSE) between the predicted state and the ground truth state, and we scale the RMSE through dividing it by the average step size for a sensible metric across dimensions following the practice in \cite{Kalman1960AOK}. Figure 1 shows the experiment results in the Maze 1, Maze 2 and Maze 3. SDPF converges to the lowest RMSE during training process. Figure 1 (a) shows the RMSE at the last time step evaluated on validation data. The SDPF exhibits the smallest RMSE while converges the slowest on validation data among SDPF, DPF and LSTM for all mazes environments. The convergence speed of RMSE for LSTM is faster but with larger RMSE than that of DPF on Maze 1 and Maze 2. Among different sizes of maze environments, SDPF converges to the lowest RMSE.

SDPF significantly improves tracking performance on testing trajectories. Figure 1 (b) shows the boxplot of RMSE at the specified time step evaluated on testing data. The SDPF has the lowest median RMSE value comparing to the DPF and LSTM for all mazes environments. Among different sizes of maze environments, SDPF possesses the lowest median RMSE. Figure 1 (c) shows one representative example of tracking performance, which provides the visualisation comparison of specified trajectory along the whole time steps on testing data in Maze 1, Maze 2 and Maze 3. It shows that SDPF has the lowest RMSE at a wide range of time steps than the other two algorithms.

SDPF is robust to wide range of percentage of labelled data. Figure 1 (d) presents tracking performance of SDPF, DPF and LSTM with different percentage of available ground truth state values during training. It shows that SDPF has the lowest RMSE value under a wide range of percentage of labelled data.
with standard deviations 30 and 1000, respectively. The initial particles are sampled from a Gaussian distribution centered at the true initial states.

To formulate the semi-supervised setting, we take trajectories with 100 time steps from 47 previously unseen buildings. To track the state of the robot under a random walk of a robot in the buildings, and the average size of the buildings and rooms are 206 m² and 37 m² respectively.

The goal of this task is to track the state of the robot under a sequence of global map, visual images, and odometry. The visual images come from a monocular RGB camera with image size of 56 × 56.

The training set is comprised of 74,800 trajectories with 25 time steps, and the validation set contains 830 trajectories with 100 time steps, and the testing set consists of 820 trajectories with 100 time steps from 47 previously unseen buildings. To formulate the semi-supervised setting, we take 1000 trajectories and randomly set the ground truth data to be unknown for training with different labelled ratio. For training and testing of DPF and SDPF, the number of particles is set to be 30 and 1000, respectively. The initial particles are sampled from a Gaussian distribution centered at the true initial states with standard deviations 0.30 m, 0.30 m, 30° for translation in x and y axes and the rotation, correspondingly. In dynamic model \( s_i = s_{i-1} + f_0(s_{i-1}, a_t) + \text{diag}(\sigma_x, \sigma_y, \sigma_\beta) \eta_i \), where \( \eta_i \sim \mathcal{N}(0; I), \sigma_x, \sigma_y, \sigma_\beta \) are set to be 0.04 m, 0.04 m, and 5°, respectively. For SDPF, we choose block length \( L = 4 \) time steps. The optimisation of parameter \( \theta \) is based on RMSProp optimiser [35] with 0.5 decay rate, and the initial learning rate is set to be 0.0001 as in [21].

2) Experimental results: We evaluate the performance of SDPF following the same procedure as in [21]. The main differences between House3D and Maze environments are: (i) The testing environments for House3D are previously unseen environments while they are the same as training environments for Maze, (ii) For the House3D, the likelihood of the predicted particles in measurement model combines features of floor map and observation image, while in the Maze, it is defined by a function of the similarity between features of the observation image and the particle state. Figure 2 shows the comparison between SDPF, DPF and LSTM on the testing set.

**SDPF can generalise to different environments.** Figure 2 (a) shows the boxplot of RMSE at specified time steps with 10% of labelled data on the testing data. The SDPF leads to the lowest median RMSE and smallest interquartile range (IQR) value. The SDPF can produce the most accurate tracking results compared with DPF and LSTM. Figure 2 (b) shows the RMSE from all trajectories in testing set with different percentage of ground truth samples during training. The SDPF is able to produce significantly lower RMSE compared to DPF especially when trained with small percentage of labelled data such as 5% and 10%. The standard error of RMSE for SDPF is consistently smaller comparing to DPF and LSTM. Figure 3 shows a visualisation of the tracking performance of SDPF in Maze 1, Maze 2, Maze 3 and House3D.

**VI. Conclusion**

We propose an end-to-end semi-supervised learning method using maximum pseudo-likelihood estimation for differentiable particle filters (SDPF), which can leverage unlabelled states and the history of observations and actions to learn the dynamic and measurement models. The proposed SDPF outperforms DPF and LSTM in state estimation tasks for both global localisation and tracking when a large portion of ground truth data are unknown.

The semi-supervised differentiable particle filter can be further extended in several directions. For example, invertible neural network can be introduced to replace the conventional neural network and enable flexible transformation while keeping both density estimation and sampling computationally tractable. Another research area is unsupervised learning and self-supervised learning for effective parameter inference in differentiable particle filters without any ground truth data.
