

Adaptive Probabilistic Vehicle Trajectory Prediction Through Physically Feasible Bayesian Recurrent Neural Network

Chen Tang, Jianyu Chen and Masayoshi Tomizuka

Abstract—Probabilistic vehicle trajectory prediction is essential for robust safety of autonomous driving. Current methods for long-term trajectory prediction cannot guarantee the physical feasibility of predicted distribution. Moreover, their models cannot adapt to the driving policy of the predicted target human driver. In this work, we propose to overcome these two shortcomings by a Bayesian recurrent neural network model consisting of Bayesian-neural-network-based policy model and known physical model of the scenario. Bayesian neural network can ensemble complicated output distribution, enabling rich family of trajectory distribution. The embedded physical model ensures feasibility of the distribution. Moreover, the adopted gradient-based training method allows direct optimization for better performance in long prediction horizon. Furthermore, a particle-filter-based parameter adaptation algorithm is designed to adapt the policy Bayesian neural network to the predicted target online. Effectiveness of the proposed methods is verified with a toy example with multi-modal stochastic feedback gain and naturalistic car following data.

I. INTRODUCTION

Safety is critical for autonomous vehicle (AV), because it needs to interact with the complicated driving environment filled with human-driven vehicles and pedestrians. It is then necessary for AV to predict other vehicles' future trajectories accurately. Many efforts have been made to enhance prediction accuracy [1]–[5]. However, deterministic prediction is not sufficient to enable robust safety because human driver's decisions are uncertain even under the same circumstances. Instead, the probabilistic distribution of trajectories should be estimated.

Various methods exist in literature for probabilistic trajectory prediction. Physics-based motion models are developed based on vehicle dynamics or kinematic model. Certain assumption on control inputs is made to obtain a self-evolving model, for instance, constant turn rate and velocity (CTRV) model [6]. Stochastic trajectory prediction can be achieved through Bayesian filtering [7] or Monte Carlo simulation [8], [9]. The predicted distribution satisfies certain physical constraints imposed by the vehicle model. However, they are precise only for short time horizon [10]. Another stream of works skips the physics and directly constructs a probabilistic model of trajectory. Commonly used models include Gaussian mixture model (GMM) [11] and Gaussian Process (GP) [12]. Moreover, the prevalence of deep learning facilitated the usage of deep recurrent neural

networks (RNN) in probabilistic trajectory prediction [13], [14]. This category of methods have better performance for longer prediction horizon but feasibility is not guaranteed. In practice, they are usually combined with maneuver recognition module, resulting in a hierarchical structure [15]. However, human driver's intention in complicated scenarios can hardly be labeled, preventing accurate training of the maneuver recognition module.

To enhance robustness, the models need to be trained with trajectories executed by different human drivers to avoid overfitting with specific driving strategies. On the other hand, when deploying the trained models, the specific target driver should have relatively consistent characteristics over time. Therefore, its policy should have smaller variance compared to the distribution in the training dataset. Consequently, accuracy could be increased if we are able to adapt the model to the specific target of interest online. This issue is considered in [16]. Individual driving characteristics are encoded as parameters in potential field functions, which are identified online. However, its performance is restricted by its relatively simple and inflexible model structure. Moreover, it is designed for a particular scenario without the capability for generalization.

In this work, we aim to combine the advantages of both categories of methods. Concretely, our objective is to design a probabilistic model predicting dynamically feasible trajectory distribution for long horizon. More importantly, we want to make it adaptable to the specific target vehicle online. We divide the entire model into two components, policy and dynamics. The policy describes how the agent reacts to his observations. The dynamics describes how the scenario propagates given his actions. Dynamics model at certain level of complexity can normally be assumed known for vehicles [17]. It can then be combined with an expressive data-driven model representing the complicated intrinsic policy, forming an unit model for a single time step. To infer trajectory distribution, units for the entire prediction horizon are then connected into a recurrent structure. In this case, the problem is transformed into a probabilistic policy learning problem, which is similar to the idea of probabilistic model-based imitation learning [18].

We propose to model the policy with Bayesian neural network (BNN) which enables complicated output distribution, e.g. multi-modal distribution. It can therefore unify the commonly used heuristic-based hierarchical model structure, so that the policy model is optimized as a whole. Potentially, it can be connected with deep-learning-based perception and decision making modules. Bayesian deep learning can

[†]This work was supported by DENSO International America, Inc

[‡]C. Tang, J. Chen and M. Tomizuka are with the Department of Mechanical Engineering, University of California, Berkeley, CA 94720 USA (e-mail: chen_tang@berkeley.edu, jianyuchen@berkeley.edu, tomizuka@berkeley.edu).

be applied in an end-to-end fashion to optimize the entire architecture [19]. Its flexibility in distribution representation also enables a policy model for general driving scenarios, where a single parametric family of distribution is not sufficient. We adopt a gradient-based method black-box α -divergence minimization (BB- α) [20] to train the entire recurrent model for better long-term prediction performance. Eventually, we obtain a Bayesian recurrent neural network (BRNN). To achieve adaptation, we propose a particle-filter-based online adaptation method to incorporate new samples quickly online. The model after training is treated as a prior to obtain the posterior policy BNN adapted to the target human driver.

Contribution of this work is threefold: 1) We propose to convert the trajectory prediction problem into a policy learning problem and develop a BRNN model for long-term trajectory prediction. (2) Particle-filter-based online adaptation method is proposed to adapt the policy BNN model online to the specific predicted target. (3) The proposed methods are applied to solve a practical vehicle trajectory prediction problem in car following scenario.

II. BACKGROUND

A. Bayesian Neural Networks with Stochastic Inputs

A BNN with random input noise z is developed in [21] for stochastic dynamical system modeling. Given data $\mathcal{D} = \{\mathbf{x}_n, \mathbf{y}_n\}_{n=1}^N$, where \mathbf{x}_n is feature vector and \mathbf{y}_n is labeled output, it is assumed that $\mathbf{y}_n = g(\mathbf{x}_n, z_n; \mathcal{W}) + \epsilon_n$, where $g(\cdot, \cdot; \mathcal{W})$ represents a neural network with L layers and L_l hidden units in layer l , parameterized by $\mathcal{W} = \{\mathbf{W}_l\}_{l=1}^L$, a collection of $V_l \times (V_{l-1} - 1)$ weight matrices. The elements of \mathbf{W}_l are mutually independent and an identical Gaussian prior is assigned to each of them, i.e. $w_{ij,l} \sim \mathcal{N}(0, \lambda)$. z_n is a synthetic stochastic disturbance to enable richer probabilistic distribution structure. Each pair of sample $(\mathbf{x}_n, \mathbf{y}_n)$ has a corresponding independent z_n . ϵ_i is random noise to enable continuous distribution of \mathbf{y}_n . Both z_n and ϵ_i have Gaussian priors, i.e. $z_n \sim \mathcal{N}(0, \gamma)$ and $\epsilon_n \sim \mathcal{N}(0, \Sigma_\epsilon)$. The model is trained by estimating the posterior distribution of parameters $p(\mathcal{W}, z|\mathcal{D})$, where z is the vector collecting all the random inputs. The distribution $p(\mathcal{W}, z|\mathcal{D})$ is approximated by $q(\mathcal{W}, z)$ as in (1). The conditional output distribution $p(\mathbf{y}|\mathbf{x}, \mathcal{D})$ can then be approximated by Monte Carlo sampling, i.e. computing samples of \mathbf{y} with samples of \mathcal{W}, z drawn from $q(\mathcal{W}, z)$. The structural complexity and non-linearity embedded in the neural network enable abundant probabilistic structure, strengthening the expressiveness of the model.

$$q(\mathcal{W}, z) = \prod_{l=1}^L \prod_{i=1}^{V_l} \prod_{j=1}^{V_{l-1}+1} \mathcal{N}(w_{ij,l} | m_{ij,l}^w, v_{ij,l}^w) \prod_{n=1}^N \mathcal{N}(z_n | m_n^z, v_n^z) \propto \left[\prod_{n=1}^N f(\mathcal{W}) f_n(z_n) \right] p(\mathcal{W}) p(z), \quad (1)$$

where

$$p(\mathcal{W}) = \prod_{l=1}^L \prod_{i=1}^{V_l} \prod_{j=1}^{V_{l-1}+1} \mathcal{N}(w_{ij,l} | 0, \lambda) \\ p(z) = \prod_{n=1}^N \mathcal{N}(z_n | 0, \gamma) \\ f_n(z_n) = \exp \left\{ \frac{v_n^z - \gamma}{2\gamma v_n^z} z_n^2 + \frac{m_n^z}{v_n^z} z_n \right\} \\ f(\mathcal{W}) = \exp \left\{ \sum_{l=1}^L \sum_{i=1}^{V_l} \sum_{j=1}^{V_{l-1}+1} \frac{1}{N} \left(\frac{v_{ij,l}^w - \lambda}{2\lambda v_{ij,l}^w} w_{ij,l}^2 + \frac{m_{ij,l}^w}{v_{ij,l}^w} w_{ij,l} \right) \right\}$$

B. Black-Box α -Divergence Minimization

BB- α [20] can be used to approximate the distribution $q(\mathcal{W}, z)$, which is a gradient-based and sampling-based approximate inference method. At each iteration, M samples of \mathcal{W} and z are drawn from current estimated posterior distribution. A mini-batch of data $\mathcal{S} \subset \{1, 2, 3, \dots, N\}$ is sampled uniformly. The noisy estimate of energy function as shown in (2) is minimized by tuning $q(\mathcal{W}, z)$ and Σ through gradient descent. The only term related to the data is the likelihood function $p(\mathbf{y}_n | \mathbf{x}_n, \mathcal{W}_s, z_{n,s}, \Sigma)$. One advantage of BB- α is that the divergence function used in the energy function can be tuned by adjusting α , providing flexibility to improve performance [20]. A general guidance for α value selection is provided in [21]. In practice, the best α value can be directly found by validation.

$$\hat{E}_\alpha(q) = -\log Z_q - \frac{N}{\alpha |\mathcal{S}|} \sum_{n \in \mathcal{S}} \hat{L}_\alpha^n(q), \quad (2)$$

where

$$\log Z_q = \sum_{l=1}^L \sum_{i=1}^{V_l} \sum_{j=1}^{V_{l-1}+1} \left(\frac{1}{2} \log(2\pi v_{ij,l}^w) + \frac{(m_{ij,l}^w)^2}{2v_{ij,l}^w} \right) \\ + \sum_{n=1}^N \left[\frac{1}{2} \log(2\pi v_n^z) + \frac{(m_n^z)^2}{2v_n^z} \right] \\ \hat{L}_\alpha^n(q) = \log \frac{1}{M} \sum_{s=1}^M \left(\frac{p(\mathbf{y}_n | \mathbf{x}_n, \mathcal{W}_s, z_{n,s}, \Sigma)}{f(\mathcal{W}_s) f(z_{n,s})} \right)^\alpha.$$

III. PROBLEM FORMULATION

In this section, we formulate the trajectory prediction problem and introduce the notations we use in latter sections. The problem is formulated in the context of the interactive car following scenario that our proposed methods are verified with. It can be generalized to other scenarios and other trajectory prediction problems easily. The following vehicle, denoted by P , is driven by human driver and the preceding vehicle, denoted by Q , is an AV. We define a state variable $\mathbf{x} = [d_{pq} \ v_p \ v_q]^T$. d_{pq} is the longitudinal distance between the two vehicles. v_p and v_q are the longitudinal speed of P and Q . The actions executed by the two vehicles are denoted as $\mathbf{a}_p = [\Delta d_p \ \Delta v_p]^T$ and $\mathbf{a}_q = [\Delta d_q \ \Delta v_q]^T$. Δd_p and Δd_q are the longitudinal displacement of P and

Q over the sampling time Δt . Δv_p and Δv_q are the change of longitudinal speed of P and Q over Δt . The scenario is modeled as a discrete-time system whose dynamics is described below and illustrated in Figure 1.

$$d_{pq,i+1} = d_{pq,i} - \Delta d_{p,i} + \Delta d_{q,i} + \omega_{1,i} \quad (3)$$

$$v_{p,i+1} = v_{p,i} + \Delta v_{p,i} + \omega_{2,i} \quad (4)$$

$$v_{q,i+1} = v_{q,i} + \Delta v_{q,i} + \omega_{3,i}. \quad (5)$$

In the equations, $\omega_{1,i}$, $\omega_{2,i}$ and $\omega_{3,i}$ are additive Gaussian noise accounting for the uncertainty in dynamics. We denote $\omega_i \sim \mathcal{N}(0, \Sigma_{\omega_i})$ as the random vector collecting these three noise variables. For convenience, the dynamics model is denoted as the following general form in the remainder:

$$\mathbf{x}_{i+1} = h(\mathbf{x}_i, \mathbf{a}_{p,i}, \mathbf{a}_{q,i}) + \omega_i \quad (6)$$

we assume we know \mathbf{x} for the current and past time steps and the planned future actions of the AV, \mathbf{a}_q , for next h time steps. At a time step k , given a trajectory of \mathbf{x} from time step $k-m$ to k , $\mathbf{x}_{k-m:k}$, and a series of \mathbf{a}_q from time step k to $k+h-1$, $\mathbf{a}_{k:k+h-1}$, the conditional distribution of future state trajectory $p(\mathbf{x}_{k+1:k+h} | \mathbf{x}_{k-m:k}, \mathbf{a}_{q,k:k+h-1})$ is predicted. The predicted distribution is utilized by sampling-based motion planning algorithms which consider interaction and uncertainty [22].

To fulfill the goal, we need to construct a probabilistic model of trajectory. In this work, we break down the probabilistic model into units connected recurrently. Each unit corresponds to a single time step, describing the transition of state at this time step. At a time step i , we assume that the state at next time step \mathbf{x}_{i+1} depends on the states of previous $m+1$ time steps, i.e. $\mathbf{x}_{i-m:i}$. Consequently, each unit describes the probabilistic transition function $p_{\theta_i}(\mathbf{x}_{i+1} | \mathbf{x}_{i-m:i}, \mathbf{a}_{q,i})$, parameterized by θ_i for $i = k, \dots, k+h-1$. By assuming the unit is time-invariant, the transition functions at all time steps are parameterized by an identical parameter vector θ . The probabilistic transition function can be further split up as follows:

$$p_{\theta}(\mathbf{x}_{i+1} | \mathbf{x}_{i-m:i}, \mathbf{a}_{q,i}) = \int p(\mathbf{x}_{i+1} | \mathbf{x}_i, \mathbf{a}_{q,i}, \mathbf{a}_{p,i}) p_{\theta}(\mathbf{a}_{p,i} | \mathbf{x}_{i-m:i}) d\mathbf{a}_{p,i}. \quad (7)$$

The function $p_{\theta}(\mathbf{a}_i | \mathbf{x}_{i-m:i})$ represents P 's inherent policy. We assume that it depends only on $\mathbf{x}_{i-m:i}$, but not on $\mathbf{a}_{q,i}$. It is reasonable as human drivers make decisions according to their current and historical observations. Other drivers' decisions cannot be directly observed. $p(\mathbf{x}_{i+1} | \mathbf{x}_i, \mathbf{a}_{p,i}, \mathbf{a}_{q,i})$ represents the dynamics. Since the dynamics model is known as (6), we can obtain $p(\mathbf{x}_{i+1} | \mathbf{x}_i, \mathbf{a}_{p,i}, \mathbf{a}_{q,i}) = \mathcal{N}(\mathbf{x}_{i+1}; h(\mathbf{x}_i, \mathbf{a}_{p,i}, \mathbf{a}_{q,i}), \Sigma_{\omega_i})$. By contrast, probabilistic model of the policy is unknown and need to be identified, which is the key element in our proposed model.

IV. METHODOLOGY

In this work, we modify the BNN with random input noise introduced in previous section to model the stochastic control policy. To enhance prediction performance for long

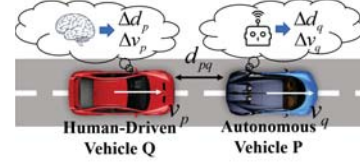


Fig. 1: Interactive Car Following Scenario with human-driven vehicle Q and autonomous vehicle P .

time horizon, the entire recurrent model consisting of policy BNN and physical model is trained via BPTT. Accordingly, the likelihood function in (2) is replaced by the likelihood of labeled future trajectory, estimated with BNN parameter sampling and optimal filtering. The parameters of the policy BNN are adapted online through particle filtering. Intuitively, weight samples generating trajectory similar to the actual trajectory are assigned higher weights. Consequently, the mass of weight distribution is shifted toward those weights.

A. Physically Feasible Bayesian Recurrent Neural Network

We use the BNN with random input noise to represent the policy. Different from [21], we assume that the random inputs for all the samples are identical, i.e. $z_1 = z_2 = \dots = z_N = z$. Otherwise, it is ambiguous which z_n should be used for a test sample. Concretely, we assume that $\mathbf{a}_{p,i} = g(\mathbf{x}_{i-m:i}, z; \mathcal{W}) + \epsilon_i$ at each time step i . The posterior distribution is then approximated with (8) instead.

With the policy BNN, the recurrent model developed in last section becomes a BRNN with dynamics model embedded (Figure 2). Instead of training the policy BNN alone, we propose to directly train the entire recurrent model. Since BB- α is a gradient-based method, it is straightforward to extend the scheme to train the BRNN via BPTT [23], [24]. The training dataset consists of state trajectories, planned future input series and actual future trajectories, i.e. $\mathcal{D} = \{\mathbf{x}_{-m:0}^n, \mathbf{a}_{q,0:h-1}^n, \mathbf{x}_{1:h}^n\}_{n=1}^N$. In addition to \mathcal{W} and z , samples of additive noise series $\{\omega_{0:h-1}^n\}_{n=1}^N = \{\omega^n\}_{n=1}^N$ and $\{\epsilon_{0:h-1}^n\}_{n=1}^N = \{\epsilon^n\}_{n=1}^N$ are drawn at each iteration. The energy function is essentially the same. The major difference is that the likelihood function becomes $p(\mathbf{x}_{1:h}^n | \mathbf{x}_{-m:0}^n, \mathbf{a}_{q,0:h-1}^n, \mathcal{W}_s, z_s, \epsilon_s^n, \omega_s^n)$, namely the likelihood of labeled future trajectory given its corresponding historical trajectory, future input series and sampled parameters. Besides, the energy function should be revised based on the assumption on identical z over samples. It is worth noting that policy BNNs for the entire time sequence share the same \mathcal{W}_s and z_s under the sampling strategy. It indicates an underlying assumption of invariant policy over the given time horizon h . It enables our particle-filter-based online parameter adaptation algorithm explained later in this section.

One problem remained is how to estimate the likelihood. We follows a simple estimation method to save computational time. Given a sample $\mathbf{x}_{i-m:i,s}^n$ at time step i and corresponding sample of neural network parameters \mathcal{W}_s and z_s , we compute the mean of predicted action $\bar{\mathbf{a}}_{p,i,s}^n$. The conditional distribution of predicted action

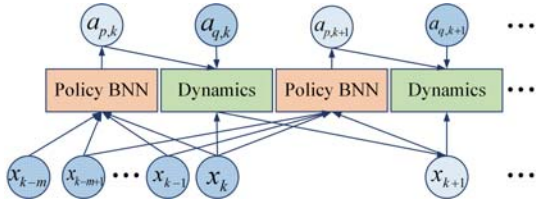


Fig. 2: Structure of the physically feasible BRNN. The shadowed nodes $\mathbf{x}_{k-m:k}$ and $\mathbf{a}_{q,k:k+h-1}$ are given.

then follows $\mathcal{N}(\bar{\mathbf{a}}_{p,i,s}^n, \Sigma)$. Since the dynamics model is known, we can then estimate the conditional distribution $p(\mathbf{x}_{i+1,s}^n | \mathcal{W}_s, z_s, \mathbf{a}_{q,i}^n, \mathbf{x}_{i-m,i,s}^n)$ as $\mathcal{N}(\mathbf{m}_{i+1,s}^n, \Sigma_{i+1,s}^n)$, following the prior update procedure of optimal filtering [25], e.g. Kalman filter (KF) for linear system and extended Kalman filter (EKF) for nonlinear system. A sample $\mathbf{x}_{i+1,s}^n$ can be computed with $\bar{\mathbf{a}}_{p,i,s}^n$ and sampled noises $(\epsilon_{i,s}^n, \omega_{i,s}^n)$, along with the dynamics model for the usage of computation for later time steps. Eventually, we can approximate the likelihood function with $\prod_{i=1}^h \mathcal{N}(\mathbf{x}_i^n | \mathbf{m}_{i+1,s}^n, \Sigma_{i+1,s}^n)$. In perspective of a single sample of neural network parameters, this estimation scheme underestimates the propagation of variance through time. However, the problem can be alleviated by increasing the number of samples J to decrease the prior weight of each sample.

$$q(\mathcal{W}, \mathbf{z}) = \prod_{l=1}^L \prod_{i=1}^{V_l} \prod_{j=1}^{V_{l-1}+1} \mathcal{N}(w_{ij,l} | m_{ij,l}^w, v_{ij,l}^w) \mathcal{N}(z | m^z, v^z)^N \propto [f(\mathcal{W})f(z)]^N p(\mathcal{W})p(z)^N, \quad (8)$$

where

$$p(z) = \mathcal{N}(z | 0, \gamma), f(z) = \exp \left\{ \frac{v^z - \gamma}{2\gamma v^z} z^2 + \frac{m^z}{v^z} z \right\}$$

Training the entire recurrent model has several benefits. Firstly, the distribution $q(\mathcal{W}, \mathbf{z})$ approximates the posterior distribution conditioned on sampled trajectories. Therefore, the approximated posterior conditional distribution of trajectory inferred from the model can resemble the actual distribution better. Moreover, since the dynamics model is embedded into the recurrent model, it confines the output distribution to probability with dynamical constraints. Consequently, the parameters are optimized over dynamically feasible family of output distribution.

B. Online Parameter Adaptation via Sequential Monte Carlo

Gradient-based methods are not suitable for real-time parameter adaptation since convergence cannot be guaranteed within certain iterations. Instead, we adopt sequential Monte Carlo (SMC) sampling, namely particle filtering. SMC has been adopted for neural network training in early stage [26]. It is computationally expensive for large-scale training. In contrast, new samples collected in real time can be incorporated into posterior distribution estimation by one-step update sequentially.

Formally, we assume the weights of the policy BNN are invariant over time. Equation (6) together with the model of

Algorithm 1 Parameter adaptation algorithm

Input: $q_{k-u}(\mathcal{W}, \mathbf{z})$, $\mathbf{x}_{k-m-u:k}$, $\mathbf{a}_{q,k-u:k-1}$
Output: new approximate posterior distribution $q_k(\mathcal{W}, \mathbf{z})$

- 1: for $s = 1, \dots, M$, draw samples from $q_{k-u}(\mathcal{W}, \mathbf{z})$ and denote them as $\{\mathcal{W}_s\}_{s=1}^M$ and $\{z_s\}_{s=1}^M$.
- 2: **for** $s = 1, \dots, M$ **do**
- 3: **for** $r = 0, \dots, u-1$ **do**
- 4: $\bar{\mathbf{a}}_{p,k-u+r,s} = g(\mathbf{x}_{k-m-u+r:k-u+r}, z_s; \mathcal{W}_s)$
- 5: approximate the likelihood by Gaussian distribution using Kalman filter or extended Kalman filter:
 $p(\mathbf{x}_{k-u+r+1} | \mathcal{W}_s, z_s, \mathbf{a}_{q,k-u+r}, \mathbf{x}_{k-m-u+r:k-u+r}) \approx \mathcal{N}(\mathbf{x}_{k-u+r+1}; \mathbf{m}_{k-u+r+1,s}, \Sigma_{k-u+r+1,s})$
- 6: **end for**
- 7: $\beta_s = \prod_{r=0}^{u-1} \mathcal{N}(\mathbf{x}_{k-u+r+1}; \mathbf{m}_{k-u+r+1,s}, \Sigma_{k-u+r+1,s})$
- 8: **end for**
- 9: estimate the mean and variance as:

$$m_{ij,l}^w = \frac{\sum_{s=1}^M \beta_s w_{ij,l,s}}{\sum_{s=1}^M \beta_s}$$

$$v_{ij,l}^w = \frac{\sum_{s=1}^M \beta_s (w_{ij,l,s} - m_{ij,l}^w)^2}{\sum_{s=1}^M \beta_s}$$

$$m^z = \frac{\sum_{s=1}^M \beta_s z_s}{\sum_{s=1}^M \beta_s}$$

$$v^z = \frac{\sum_{s=1}^M \beta_s (z_s - m^z)^2}{\sum_{s=1}^M \beta_s}$$

10: obtain $q_k(\mathcal{W}, \mathbf{z})$ based on (8)

11: **return** $q_k(\mathcal{W}, \mathbf{z})$

policy BNN are treated as measurement model of \mathcal{W} and z . An additional random noise $v_i \sim \mathcal{N}(0, \Sigma_v)$ is added in (6). The collected \mathbf{x}_{i+1} is then regarded as a noisy measurement of state, so that the parameters do not overfit with the new samples. To further prevent overfitting, parameters are updated each u steps. At each update step, the posterior distribution of BNN parameters is estimated using particle filter. The update procedure is described in Algorithm 1. For faster resampling, the weighted samples of \mathcal{W} and z are fitted to a Gaussian distribution, from which new samples are drawn [27].

V. EXAMPLES

A. Toy Linear System with Stochastic Feedback Gain

We start with evaluating the proposed method with a toy example, so that the predicted conditional distribution of trajectory can be compared directly with the actual distribution. The designed toy example is a linear system with stochastic feedback gain described in Equations (9-10).

$$x_{i+1} = x_i - \gamma S(\kappa_i) x_i \quad (9)$$

$$\kappa_{i+1} = \kappa_i + \zeta_i, \quad (10)$$

where $\kappa_0 \sim p_1 \mathcal{N}(\mu_1, v_1) + p_2 \mathcal{N}(\mu_2, v_2)$ and $\zeta_i \sim \mathcal{N}(0, v_\zeta)$. The function $S(\cdot)$ is the sigmoid function. Along with the constant γ , it enables stability of the closed-loop system. A random noise ζ_i is added to have κ_i vary over time. Given an initial state x_0 , we want to enquire the conditional distribution $p(x_{1:15} | x_0)$ from the model. The policy BNN is

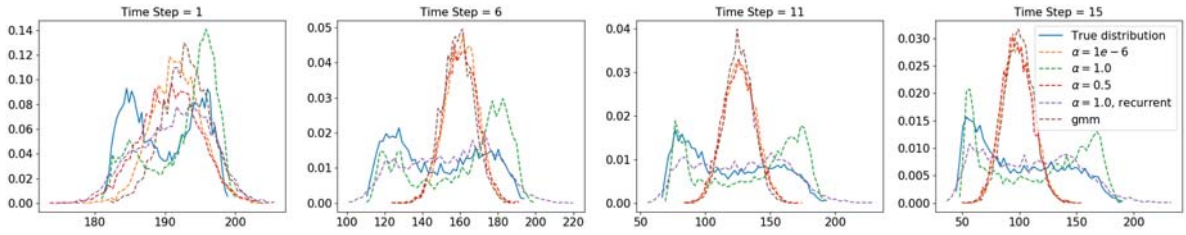


Fig. 3: Normalized histograms of x_k at different time steps. The solid blue line represents histogram of actual x_k , dash lines with different color represent histogram of x_k sampled from different models. The x-axis is the value of x_k and the y-axis is the corresponding approximated density.

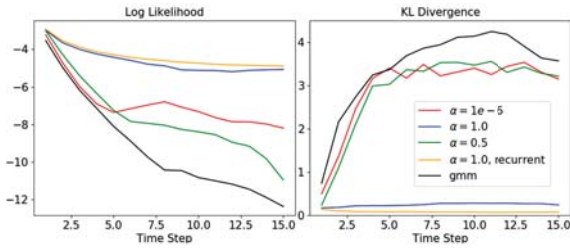


Fig. 4: Estimated log likelihood and KL-divergence of x_k over the 15 time steps. The log likelihood is estimated as in the training stage using 100 sampled trajectories. The KL Divergence $KL(p_{\text{actual}}||p_{\text{approx}})$ is estimated based on the normalized histograms, where p_{actual} is the actual distribution and p_{approx} is the predicted distribution.

expected to approximate the average gain distribution over the prediction horizon.

We set $x_0 = 200$, $\mu_1 = -1$, $\mu_2 = 1$, $v_1 = v_2 = 0.36$, $p_1 = p_2 = 0.5$ and $v_\zeta = 0.04$. 5000 samples of $x_{1:15}$ are drawn for training. The network has two hidden layers with 50 hidden units per layer. The batch size is 50. The learning rate is chosen as $1e - 4$. 100 samples are drawn to estimate the energy function. The model is implemented in Tensorflow [28]. We compare BNNs trained with three special α values, $1e - 6$, 0.5 and 1.0. Afterwards, BRNN is trained with the best α value found. Moreover, a conditional GMM is trained by Expectation Maximization (EM) algorithm for comparison, so that the benefit of the complicated probability structure enabled by BNN could be illustrated. The GMM model implemented in scikit-learn [29] is adopted. 10 mixture components are selected after tuning. BNN with $\alpha = 1.0$ captures the multi-modal structure of distribution, while as BNNs with other α values as well as GMM model fail to do so. Its performance surmounts other models in terms of log likelihood and KL-divergence as well. Training the entire recurrent model further increases the log likelihood and decreases the KL-divergence, especially for latter time steps in the prediction horizon. BRNN tends to approximate the state distribution better over the entire prediction horizon, at the cost of relatively lower similarity at the first time step. The online parameter adaptation algorithm is tested by adapting the policy BNN to trajectories generated with $\kappa_0 \sim \mathcal{N}(\mu_1, v_1)$. 9 adaptation iterations are applied. For each iteration i , the policy BNN is adapted to a sampled

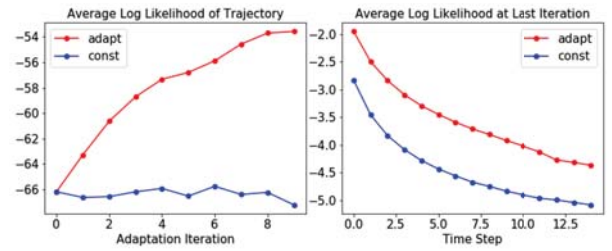


Fig. 5: (a) Average log likelihood over adaptation iterations. (b) Average log likelihood over the prediction horizon. Adapted model after the last iteration is compared with the original model.

trajectory $x_{0:30}^i$. Log likelihood of the trajectory $x_{0:15}^i$ under the predicted distribution is recorded. Meanwhile, the log likelihood computed with the unadapted model is also recorded for comparison. The test procedure is conducted for 50 times. The average log likelihood increases over iterations (Figure 5), indicating adaptation towards the distribution of new samples.

B. Interactive Car Following Model from Traffic Data

To verify the proposed method for stochastic trajectory prediction problem in autonomous driving, we start with learning an interactive car following model from naturalistic traffic data. The scenario is stated in Section III. The dataset selected is Interstate 80 (I-80) Freeway Dataset from the Next Generation Simulation (NGSIM) program [30]. We set $\Delta t = 0.2s$, $m = 9$ and $h = 30$. The network has three hidden layers with 50 hidden units per layer. The batch size is 250. The learning rate is chosen as $1e - 4$. 1000 samples are drawn to estimate the energy function. During the training of the policy BNN, $\alpha = 0.5$ achieves the best performance in terms of log likelihood. Therefore, we fix $\alpha = 0.5$ and evaluate the effect of trajectory length in the training stage of the recurrent model. Particularly, we set $h = 1$, $h = 10$, $h = 20$ and $h = 30$ in (2) and train four models correspondingly. Meanwhile, similar to the toy example, we train a GMM policy unit for comparison. 20 mixture components are selected.

We test the trained models with 100 randomly selected trajectories from test dataset. To compare the performance in short and long time horizons separately, the average log likelihood of the first 15 time steps and the one of the

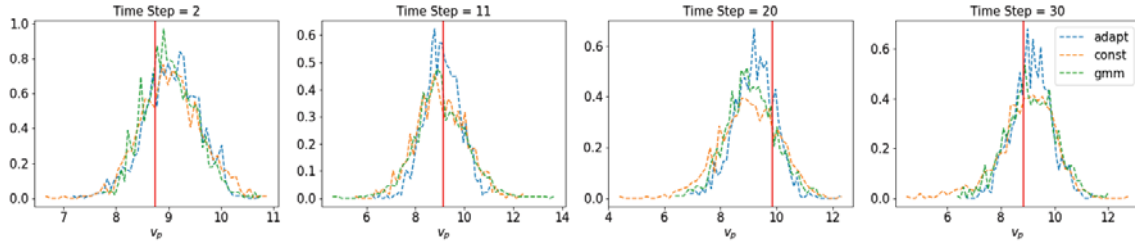


Fig. 6: Normalized histograms of predicted v_p at different time steps. The solid red vertical line represents the actual value in the dataset. The x-axis is the value of v_p and the y-axis is the corresponding approximated density.

latter 15 time steps are computed respectively. Increasing h in training leads to higher likelihood. Since the car following scenario is relatively simple with small variance, GMM actually performs quite well especially for short time horizon. But still, BRNN with $h = 30$ achieves higher log likelihood for latter time steps than GMM, indicating the effectiveness of training the entire BRNN in long-term trajectory prediction.

To test the online adaptation algorithm, we adapt the policy BNN to trajectories of the same vehicle over 90 time steps. The parameters are updated per 30 steps, i.e. 2 update iterations are applied for each trajectory. 44 trajectories are extracted from the test dataset for evaluation. Table I shows the average log likelihood of the trajectories over the last 30 time steps. BRNN with $h = 30$ achieves the best performance after adaptation, whereas smaller h led to decreased log likelihood. To visualize the distribution for comparison, we choose a trajectory and plot the normalized histogram of predicted v_p at different time steps for three models: the BRNN with $h = 30$ without adaptation, the BRNN with $h = 30$ after adaptation and GMM model. The variance of BRNN's prediction becomes smaller than the GMM model after adaptation, leading to larger likelihood of the actual v_p . It is difficult to rigorously analyze the cause as the actual distribution of data can hardly be extracted. Intuitively, the BRNN with $h = 30$ learns a policy invariant over the prediction horizon, so that the adaptation algorithm based on the assumption of time-invariant policy works well.

VI. DISCUSSION AND FUTURE WORK

As shown in last section, the proposed BRNN architecture achieves good performance in trajectory prediction especially for long prediction horizon. Its advantages are more apparent in the toy example, where the multi-modal policy leads to complicated distribution of trajectory. The distribution in

the car following scenario is relatively simple, therefore, GMM has a comparable performance with BRNN. However, BRNN still surmounts GMM in later time steps of the prediction horizon. In future research, the proposed method will be evaluated in driving scenarios where the distribution of trajectories is more complicated.

For the online adaptation algorithm, it has been illustrated in the toy example that the policy BNN can indeed adapt to online collected samples, when the assumption that the policy varies slightly over time holds. In real-world scenarios, the assumption might not hold. Based on the experimental results for car following, learning an time-invariant policy for long horizon during training seems to enable better adaptation effect, but its performance cannot be asserted. The algorithm will be further evaluated when the time-invariant assumption does not hold. Another limitation is a large number of samples has to be drawn to approximate the distribution of weights accurately when the size of BNN is large. In the future, we will investigate on reducing the number of adaptable parameters for online adaptation.

VII. CONCLUSION

In this work, a novel stochastic vehicle trajectory prediction method based on BNN is developed. Combining BNN-based policy model with the known physical model recurrently, the proposed BRNN architecture can generate physically feasible trajectory distribution. Gradient-based training method enables direct optimization towards better long-term prediction performance. Furthermore, a particle-filter-based parameter adaptation algorithm is designed to adapt the policy BNN towards the driving policy or predicted target. The proposed methods achieve good performance in both the toy example with multi-modal policy and real-world car following scenario.

TABLE I: Model Performance Comparison for Car Following Model Learning

| Model | Testing Data | | Without Adaptation | | After Adaptation | |
|----------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| | First 15 time steps | Second 15 time steps | First 15 time steps | Second 15 time steps | First 15 time steps | Second 15 time steps |
| BRNN, $h = 1$ | -3.27 ± 13.51 | -4.95 ± 4.76 | -4.71 ± 4.46 | -7.34 ± 2.39 | -191.13 ± 100.82 | -40.90 ± 45.52 |
| BRNN, $h = 10$ | -2.30 ± 5.50 | -3.43 ± 1.15 | -2.00 ± 1.05 | -3.83 ± 0.52 | -3.25 ± 2.04 | -6.20 ± 1.31 |
| BRNN, $h = 20$ | -2.13 ± 2.78 | -3.45 ± 0.90 | -2.08 ± 0.82 | -3.52 ± 0.22 | -2.10 ± 0.81 | -3.56 ± 0.32 |
| BRNN, $h = 30$ | -2.09 ± 2.03 | -3.39 ± 0.86 | -2.00 ± 0.73 | -3.55 ± 0.32 | -1.94 ± 0.70 | -3.50 ± 0.27 |
| GMM | -1.89 ± 2.59 | -3.46 ± 0.96 | -1.89 ± 1.03 | -3.69 ± 0.45 | - | - |

¹ In each cell, the average log likelihood of trajectory divided by the number of time steps is presented in form of mean \pm std.

² h is the number of future states considered at the training stage of BRNN.

REFERENCES

- [1] S. Lefèvre, D. Vasquez, and C. Laugier, "A survey on motion prediction and risk assessment for intelligent vehicles," *Robomech Journal*, vol. 1, no. 1, p. 1, 2014.
- [2] A. Houenou, P. Bonnifait, V. Cherfaoui, and W. Yao, "Vehicle trajectory prediction based on motion model and maneuver recognition," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pp. 4363–4369, IEEE, 2013.
- [3] C. Liu and M. Tomizuka, "Enabling safe freeway driving for automated vehicles," in *American Control Conference (ACC), 2016*, pp. 3461–3467, IEEE, 2016.
- [4] L. Lin, S. Gong, T. Li, and S. Peeta, "Deep learning-based human-driven vehicle trajectory prediction and its application for platoon control of connected and autonomous vehicles," in *The Autonomous Vehicles Symposium*, vol. 2018, 2018.
- [5] D. Jeong, M. Baek, and S.-S. Lee, "Long-term prediction of vehicle trajectory based on a deep neural network," in *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 725–727, IEEE, 2017.
- [6] A. Polychronopoulos, M. Tsogas, A. J. Amditis, and L. Andreone, "Sensor fusion for predicting vehicles' path for collision avoidance systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 3, pp. 549–562, 2007.
- [7] H. Veeraraghavan, N. Papanikolopoulos, and P. Schrater, "Deterministic sampling-based switching kalman filtering for vehicle tracking," in *Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE*, pp. 1340–1345, IEEE, 2006.
- [8] A. Eidehall and L. Petersson, "Statistical threat assessment for general road scenes using monte carlo sampling," *IEEE Transactions on intelligent transportation systems*, vol. 9, no. 1, pp. 137–147, 2008.
- [9] M. Althoff and A. Mergel, "Comparison of markov chain abstraction and monte carlo simulation for the safety assessment of autonomous cars," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1237–1247, 2011.
- [10] M. G. Ortiz, J. Fritsch, F. Kummert, and A. Gepperth, "Behavior prediction at multiple time-scales in inner-city scenarios," in *2011 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1068–1073, IEEE, 2011.
- [11] J. Wiest, M. Höffken, U. Kreßel, and K. Dietmayer, "Probabilistic trajectory prediction with gaussian mixture models," in *Intelligent Vehicles Symposium (IV), 2012 IEEE*, pp. 141–146, IEEE, 2012.
- [12] J. Joseph, F. Doshi-Velez, A. S. Huang, and N. Roy, "A bayesian nonparametric approach to modeling motion patterns," *Autonomous Robots*, vol. 31, no. 4, p. 383, 2011.
- [13] B. Kim, C. M. Kang, S. H. Lee, H. Chae, J. Kim, C. C. Chung, and J. W. Choi, "Probabilistic vehicle trajectory prediction over occupancy grid map via recurrent neural network," *arXiv preprint arXiv:1704.07049*, 2017.
- [14] S. H. Park, B. Kim, C. M. Kang, C. C. Chung, and J. W. Choi, "Sequence-to-sequence prediction of vehicle trajectory via lstm encoder-decoder architecture," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1672–1678, IEEE, 2018.
- [15] G. Xie, H. Gao, L. Qian, B. Huang, K. Li, and J. Wang, "Vehicle trajectory prediction by integrating physics-and maneuver-based approaches using interactive multiple models," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5999–6008, 2018.
- [16] H. Woo, Y. Ji, Y. Tamura, Y. Kuroda, T. Sugano, Y. Yamamoto, A. Yamashita, and H. Asama, "Trajectory prediction of surrounding vehicles considering individual driving characteristics," *International Journal of Automotive Engineering*, vol. 9, no. 4, pp. 282–288, 2018.
- [17] R. Schubert, E. Richter, and G. Wanielik, "Comparison and evaluation of advanced motion models for vehicle tracking," in *Information Fusion, 2008 11th International Conference on*, pp. 1–6, IEEE, 2008.
- [18] P. Englert, A. Paraschos, M. P. Deisenroth, and J. Peters, "Probabilistic model-based imitation learning," *Adaptive Behavior*, vol. 21, no. 5, pp. 388–403, 2013.
- [19] R. McAllister, Y. Gal, A. Kendall, M. Van Der Wilk, A. Shah, R. Cipolla, and A. V. Weller, "Concrete problems for autonomous vehicle safety: Advantages of bayesian deep learning," *International Joint Conferences on Artificial Intelligence, Inc.*, 2017.
- [20] J. M. Hernández-Lobato, Y. Li, M. Rowland, D. Hernández-Lobato, T. Bui, and R. E. Turner, "Black-box α -divergence minimization," 2016.
- [21] S. Depeweg, J. M. Hernández-Lobato, F. Doshi-Velez, and S. Udluft, "Learning and policy search in stochastic dynamical systems with bayesian neural networks," *arXiv preprint arXiv:1605.07127*, 2016.
- [22] J. Chen, C. Tang, L. Xin, S. E. Li, and M. Tomizuka, "Continuous decision making for on-road autonomous driving under uncertain and interactive environments," in *Intelligent Vehicles Symposium (IV), 2018 IEEE, IEEE*, 2018.
- [23] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.
- [24] M. Fortunato, C. Blundell, and O. Vinyals, "Bayesian recurrent neural networks," *arXiv preprint arXiv:1704.02798*, 2017.
- [25] B. D. Anderson and J. B. Moore, "Optimal filtering," *Englewood Cliffs*, vol. 21, pp. 22–95, 1979.
- [26] N. De Freitas, C. Andrieu, P. Højten-Sørensen, M. Niranjana, and A. Gee, "Sequential monte carlo methods for neural networks," in *Sequential Monte Carlo methods in practice*, pp. 359–379, Springer, 2001.
- [27] J. H. Kotecha and P. M. Djuric, "Gaussian sum particle filtering," *IEEE Transactions on signal processing*, vol. 51, no. 10, pp. 2602–2612, 2003.
- [28] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, pp. 265–283, 2016.
- [29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.
- [30] V. Alexiadis, J. Colyar, J. Halkias, R. Hranac, and G. McHale, "The next generation simulation program," *Institute of Transportation Engineers. ITE Journal*, vol. 74, no. 8, p. 22, 2004.