2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Stein Movement Primitives for Adaptive Multi-Modal Trajectory Generation

Zeya Yin[1], Tin Lai[1], Subhan Khan[2], Jayadeep Jacob[1], Yonghui Li[2], Fabio Ramos[1,3]

***Abstract*— Probabilistic Movement Primitives (ProMPs) and their variants are powerful methods for enabling robots to learn complex tasks from human demonstrations, where motion trajectories are represented as stochastic processes with Gaussian assumptions. However, despite their computational efficiency, these methods have limited expressiveness in capturing the diversity found in human demonstrations, which are typically characterized by the multi-modality of motions. For example, when picking up an object partially obscured by an obstacle, some individuals may opt to go to the right, while others may choose the left side of the object. In this paper, we introduce Stein Movement Primitives (SMPs), a novel approach to probabilistic movement primitives. We formulate motion primitive adaptation as a non-parametric probabilistic inference using Stein Variational Gradient Descent (SVGD), thus avoiding any explicit posterior distribution assumptions and enabling the direct representation of the multi-modality in human demonstrations. We illustrate how our method can adapt robot motion to different scenarios while maintaining high similarity to the original demonstrations, even when the demonstrations are multi-modal. Experimentally, we demonstrate our approach to several domain adaptation problems using the LASA dataset and with a real robotic arm.**

I. Introduction

Learning from Demonstrations (LfD) for intelligent robots has been developed over decades, as many complex daily tasks are challenging to accomplish solely through trajectory planning methods. Robots are expected to learn human-demonstrated skills and subsequently execute complex tasks or even perform new actions in varied scenarios without explicit instructions. To address this, various methods have been developed.

Building upon the benefits of Movement Primitives (MPs), Dynamic Movement Primitives (DMPs) have been developed to learn a deterministic average from provided demonstrations [1]. Additionally, trajectory representation can be achieved through probabilistic modeling, such as with probabilistic movement primitives (ProMPs) [2]. The primary advantage of these probabilistic approaches lies in their ability to represent multiple demonstrated trajectories as distributions characterized by parameters such as mean and covariance rather than a single trajectory. ProMPs accomplish this by representing the distribution over trajectories as a linear combination of a stochastic weight vector and a set of basis functions.

[1]The School of Computer Science, The University of Sydney, Australia. Email: {zyin0757, tinlai, jayadeep, fabio.ramos}@sydney.edu.au

[2]The School of Electrical Engineering and Computer Engineering, The University of Sydney, Australia. Email: {subhan.khan, yonghui.li}@sydney.edu.au

[3]NVIDIA, USA

Corresponding author: zyin0757@uni.sydney.edu.au

Given the vast array of potential tasks, it is impractical to demonstrate every possible variation to the robot. Therefore, a crucial aspect of any movement primitive framework is its ability to adapt primitives to unseen scenarios without requiring additional demonstrations. For adaptation, probabilistic approaches leverage the information encoded in the covariance to adjust the trajectory only at points of interest while mainly preserving the rest of the trajectory. Additionally, by representing the trajectory as a distribution, the probabilistic approach accounts for safety and robustness in unseen scenarios. In our work, the adaptation of movement primitives is formulated as unconstrained optimization within a probabilistic framework. Furthermore, we establish an appropriate prior to encoding both single-modal and multi-modal trajectories from demonstrations, reducing the need for additional cost terms such as trajectory smoothness and jerkiness.

In this paper, we leverage variational inference with ProMPs to tackle the challenge of learning skills from a set of multi-modal demonstrations with non-parametric distributions. We propose Stein Movement Primitives (SMPs)—a multi-modal probabilistic motion primitive framework that enables adaptation to new scenarios due to obstacles or different goals while maintaining high similarity to demonstrations. The main contributions of the paper are as follows:

- A SMP framework that can adapt trajectories not limited to Gaussian distribution via probabilistic inference while taking into account prior demonstrations, even when the demonstrations are multi-modal.
- A new differentiable prior distribution for trajectories—the smooth box prior—that acts as a soft constraint, ensuring that the starting and ending points of reproduced trajectories are constrained within predefined regions.
- Evaluations of SMPs in new environments with the presence of obstacles and new targets, in simulation and in real experiments with a robotic manipulator.

II. Related Work

There is a considerable amount of research work on adapting the movement primitives to new environments or specific tasks. Given that the presentation of demonstrations is deterministic, the DMP formulation has been expanded to mitigate obstacle interference. [3]–[6]. A recent advancement in generalized imitation learning involves the utilization of a diffeomorphism transformer to convert imitated trajectories generated by a gradient dynamical system into collision-free paths adapted to new target points within a novel scenario [7]. However, this approach overlooks trajectory probability

considerations and lacks the capability to manage multi-modal demonstrations. Within a probabilistic framework, it is possible to combine multiple primitives into one primitive to address multi-modality [8]–[10]. Alternatively, extensions to the ProMPs framework for obstacle avoidance have been investigated [11]–[13]. Unfortunately, many of them optimise deterministic trajectories and lose the probabilistic description of the primitive. Another ProMPs-based trajectory optimization method is described in [14]. However, it remains confined to the Gaussian assumption and only considers updating the mean of the trajectory parameter distribution by averaging the optimised trajectory samples. Bayesian inference emerges as a suitable method to adapt trajectories while upholding the probabilistic characterization of the primitives. It can be combined with variational inference (VI) that approximates the target distribution by gradient-based optimization. Unfortunately, typical VI methods make assumptions about the posterior distribution, such as a Gaussian posterior, which can lead to mode collapse in multi-modal problems. The most closely related work to ours, presented in [15], employs VI as a method for posterior regulation. In this approach, the adapted posterior distribution is designed to maintain the similarity of the original ProMPs while avoiding obstacles. However, both the prior and posterior candidate distributions are constrained to be simple Gaussian distributions to ensure an analytical solution. Another strategy is to model the set of trajectories as a Gaussian Mixture Model (GMM). This was used in [16] for policy imitation but did not include strategies for adaptation, for example, to avoid collisions. In [17], [18] the average moment of each component of the GMM is computed to form a Gaussian prior. Following that, the candidate target distribution is selected as a Gaussian distribution to facilitate the KL divergence computation, losing the multi-modality nature of the problem. A general-purpose Bayesian inference algorithm called Stein Variational Gradient Descent (SVGD) was developed in [19]. Extensions of SVGD have been proposed recently for planning and control [20]–[22]. In these works, trajectories are represented as particles and can capture multi-modal distributions. Differentiable cost functions can be added for collision avoidance, and trajectories are optimized by SVGD to approximate the target distribution. In this paper, we develop SVGD for movement primitives and propose adaptation strategies for collision avoidance and the incorporation of new target goal poses within a non-parametric inference procedure.

III. Preliminary

A. Probabilistic Movement Primitives

ProMPs and its variants are a class of methods for learning MPs within in a probabilistic framework. Assume we are given a set of N trajectories with position observations $\xi = \{\xi^j\}_{j=1}^N$ under the i.i.d. assumption, the single trajectory $\xi^j \in \mathbb{R}^{T\times D}$ consists of the robot's joint or Cartesian space coordinates such that $\xi^j = [y_0, y_1, ..., y_T]^j$ in a time sequence with length $t \in [0, T]$, where D is the dimension of the space. Each robot's joint or Cartesian space coordinate y_t in a trajectory ξ^j can be presented as a distribution by a linear model through a dot product of squared exponential radial basis functions (RBF) Φ_t and its corresponding weight vector w^j, with additive zero-mean Gaussian noise and variance Σ_y:

$$y_t^j = \Phi_t w^j + v, \quad v \sim \mathcal{N}(0, \Sigma_y), \tag{1}$$

where $\Phi_t \in \mathbb{R}^K$ denotes radial basis functions and is given as $\Phi_t = [\Phi(t, t_1^{'}), ..., \Phi(t, t_K^{'})]$, and $w \in \mathbb{R}^{K\times D}$ is a random variable weight matrix. Thus, y_t is Gaussian distributed in D dimensions. Given an initial Gaussian distribution of $p(w) = \mathcal{N}(w; \theta)$, where $\theta = \{\mu_w, \Sigma_w\}$, the distribution of the trajectory can be derived by marginalization as :

$$p(\xi|\theta, \Sigma_y) = \int \mathcal{N}(w; \theta) \prod_0^T \mathcal{N}(y_t; \Phi_t w, \Sigma_y) dw. \tag{2}$$

The prior $p(w)$ can be learned using the Expectation-Maximization (EM) method. During the E-step, the individual posterior of a single trajectory parameter $p(w^j|\xi^j) \propto p(\xi^j|w^j)p(w^j)$ is computed. Then, in the M-step, the updated $p(w) = \mathcal{N}(w; \theta^*)$ is derived by matching the moments of $\{p(w^j|\xi^j, \theta)\}_{j=1}^N$. Note that, in ProMPs, both the prior and likelihood terms are modelled as Gaussian distributions to obtain a closed-form solution. However, this may not be ideal in many cases, particularly in scenarios with obstacles that lead to multi-modal trajectories. Specifically, if the demonstrations are multi-modal, ProMPs may fail to represent the trajectories with the correct modality.

B. Stein Variational Gradient Descent

Variational inference (VI) is a standard procedure for Bayesian inference that approximates the target posterior distribution. It can be solved as an optimisation problem where a candidate distribution $q(x)$ is selected from a distribution family Q. During the optimisation, the KL-divergence between the target distribution $p(x)$ and the candidate distribution $q(x)$ is optimised. Here, x represents the random variable of distribution. The optimal candidate distribution $q^*(x)$ can be expressed by

$$q^*(x) = \arg\min_{q\in Q} D_{KL}(q(x)||p(x)). \tag{3}$$

However, traditional VI requires choosing the type of candidate distribution $q(x)$ to approximate the target distribution $p(x)$. Additionally, the explicit closed-form solution can only be derived in special cases, such as both $q(x)$ and $p(x)$ are Gaussian distributions. In contrast, SVGD circumvents such limitations by representing the variational distribution q with a set of particles $\{x^i\}_{i=1}^m$ [19]. These particles form an empirical distribution in the form of $\frac{1}{m}\sum_{j=i}^m \delta(x - x^i)$, where $\delta(\cdot)$ is the Dirac delta function. In each iteration of SVGD, these particles are updated for a given step size $\epsilon > 0$ in parallel: $x^i = x^i + \epsilon\phi(x)$, where $\phi(\cdot)$ is the score function. Specifically, ϕ lies in the unit-ball of a Reproducing Kernel Hilbert Space (RKHS) and defines a velocity field that minimises the KL-divergence efficiently:

$$\phi^* = \arg\max_{\phi\in\mathcal{H}} \{\nabla_\epsilon D_{KL}(q_{[\epsilon\phi]}||p(x)) : \ \|\phi\|_{\mathcal{H}} \leq 1\} \tag{4}$$

Notably, there exists a closed-form solution for ϕ^* which is interpreted as a functional gradient in RKHS,

$$\phi^*(x) = \frac{1}{m}\sum_{i=1}^{m}[k(x^i, x)\nabla_{x^i}\log p(x^i) + \nabla_{x^i}k(x^i, x)]. \quad (5)$$

In the above, $k(\cdot)$ is the kernel specifying the RKHS. In this paper, we use the squared exponential kernel. After optimization, the particles x act as empirical samples to approximate the target distribution of $p(x)$.

IV. Methodology

In this section, we introduce Stein Movement Primitives (SMPs), combining ProMPs with Stein Variational Gradient Descent (SVGD) to enable multi-modal trajectory inference. Our approach generalizes from single-modal to multi-modal trajectories, allowing adaptation to unseen scenarios and novel goal targets while maintaining a non-parametric trajectory distribution. Additionally, we incorporate a smooth box prior into the SMP framework, constraining the trajectories to specific points of interest during inference. The algorithm's workflow is outlined in Algorithm 1.

A. Stein Movement Primitives as Non-Parametric Probabilistic Inference

First, we introduce a binary optimality criterion, denoted by the random variable $\mathcal{O}$, where $\mathcal{O} \in \{0, 1\}$. The variable $\mathcal{O}$ indicates how well the reproduced trajectories generalize to new scenarios and match the demonstrations in terms of their likelihood, defined as $p(\mathcal{O} \mid w)$. Specifically, $p(\mathcal{O} = 1 \mid w)$ represents the probability that all reproduced trajectories generated by the particles w are optimal during adaptation to a new task. For brevity, we denote $\mathcal{O} = 1$ as $\mathcal{O}$ throughout the rest of the paper. This leads to the following likelihood function:

$$p(\mathcal{O}|w) \propto \exp(-\lambda\mathcal{C}(w)). \quad (6)$$

The empirical distribution of reproduced trajectories generated by particles w can be expressed as $\{\Phi w^i\}_{i=1}^m$, where $\Phi \in \mathbb{R}^{T\times K}$ is a matrix composed of Φ_t. Each coordinate of the reproduced trajectories at time t is given by $\{\Phi_t w^i\}_{i=1}^m$. Here, $\mathcal{C}(\cdot)$ denotes the cost function associated with these reproduced trajectories, which may encompass various cost terms, such as obstacle collision costs $\mathcal{C}_{\text{obs}}(w)$, modulation costs $\mathcal{C}_{\text{mod}}(w)$, or a summation of multiple cost likelihoods $\{p_g(\mathcal{O} \mid w)\}_{g=1}^G$, expressed as:

$$\begin{aligned} p(\mathcal{O}|w) &\propto \prod_{g=1}^{G} p_g(\mathcal{O}|w) \\ &= \exp(\sum_{g=1}^{G} -\lambda_g\mathcal{C}_g(w)). \end{aligned} \quad (7)$$

The parameter λ serves as a temperature hyper-parameter that balances the influence of the loss and the prior term. In subsequent sections, we will utilize $\mathcal{C}$ instead of $\mathcal{C}(w)$ for brevity and present the general meaning of costs for single tasks or a summation of multiple tasks. By applying Bayes' Rule, we can obtain the posterior distribution,

$$p(w|\mathcal{O}) \propto p(\mathcal{O}|w)p(w). \quad (8)$$

The prior $p(w)$ can be modelled as a Gaussian distribution as introduced in Sec.III-A or Gaussian Mixture Model (GMM). The posterior distribution $p(w|\mathcal{O})$ can effectively capture the probability of parameter w conditioned on the optimality criterion $\mathcal{O}$ and encode crucial information during inference. Subsequently, we employ this posterior distribution $p(w|\mathcal{O})$ as our target distribution in variational inference,

$$q^*(w) = \arg\min_{q\in Q} D_{KL}(q(w)||p(w|\mathcal{O})). \quad (9)$$

The solution that maximises the Evidence Lower Bound (ELBO) of Eq.(9) is expressed as follows:

$$q^*(w) = \arg\min_{q\in Q} -\mathbb{E}_q[p(\mathcal{O}|w)] + D_{KL}(q(w)||p(w)). \quad (10)$$

We aim to approximate the posterior distribution $p(w|\mathcal{O})$ by using a set of particles $\{w^j\}_{j=1}^m$ instead of choosing an explicit distribution. During inference at the initial time $t = 0$, the m particles are sampled from the prior distribution $p(w)$. Each particle is then updated iterative in parallel: $w^i = w^i + \epsilon\phi^*(w^i)$, where :

$$\phi^*(w) = \frac{1}{m}\sum_{i=1}^{m}[k(w^i, w)\nabla_{w^i}\log p(w^i|\mathcal{O}) + \nabla_{w^i}k(w^i, w)]. \quad (11)$$

Then, the approximated posterior distribution $p(w|\mathcal{O})$ is updated.

B. Adaptation of Stein Movement Primitives

1) Likelihood Model for Collision Avoidance: In our paper, we model a fully connected neural network as an occupancy mapping to compute the collision costs, denoted as $f_{\text{map}}(\cdot)$. The neural network has a sigmoid function in the output layer and a tanh activation function between each hidden layer, enabling it to learn a mapping between the coordinates of trajectories and their probability of occupancy. Given the reproduced trajectories $\{(\Phi_t w)\}_{t=0}^T$, consisting of T pairs of coordinates generated by SMPs and the binary value signifying the occupancy status of each coordinate, where $z_t \in \{0, 1\}$, the collision cost is expressed as $f_{\text{map}}(\Phi_t w) = P(z = 1 \mid \Phi_t w)$, where $z = 1$ represents occupancy. We can represent the collision values for one trajectory in a matrix $V_{\text{map}} = \{f_{\text{map}}(\Phi_t w)\}_{t=0}^T$, where $V_{\text{map}} \in \mathbb{R}^{T\times 1}$. The collision costs generated by the continuous occupancy map can be represented by a quadratic term, resulting in a scaled inner product: $\mathcal{C}_{\text{obs}} = \|V_{\text{map}}\|^2$. Subsequently, we can define the likelihood of a state being collision-free as $\exp(-\lambda_{\text{obs}}\mathcal{C}_{\text{obs}})$ and incorporate it into Eq. (10),

$$\begin{aligned} p_{\text{obs}}(\mathcal{O}|w) &\propto \prod_{t=0}^{T} \exp(-\lambda_{\text{obs}}\|f_{\text{map}}(\Phi_t w)\|^2) \\ &= \exp(-\lambda_{\text{obs}}\|V_{\text{map}}\|^2) \\ &= \exp(-\lambda_{\text{obs}}\mathcal{C}_{\text{obs}}). \end{aligned} \quad (12)$$

This allows us to recast the collision avoidance problem as probabilistic inference with $p_{\text{obs}}(w|\mathcal{O}) \propto p_{\text{obs}}(\mathcal{O}|w)p(w)$. Our method has a distinctive advantage: it accommodates various homotopy classes within the likelihood, a feature that arises due to the presence of obstacles. Further results on this aspect will be presented in Sec.V.

Algorithm 1: SMPs for Adaptation with Smooth Box Prior

Initialization: $p(w)$ fitted from a set of N trajectories as observation $\xi = \{\xi^j\}_{j=1:N}$
Sample $\{w^i\}_{i=1:m} \propto p(w)$
while *Not Completed* **do**
  compute posterior consisting of all particles $\{w^i\}_{i=1}^m$ for $p(w^i|\mathcal{O}) \propto \exp(-\lambda\mathcal{C})p(w^i)$
  for *i = 1,2 ...m* **do**
    $\nabla_{w^i} \log p(w^i|\mathcal{O}) \leftarrow \nabla_{w^i}(-\lambda\mathcal{C}) + \nabla_{w^i} \log p_{\text{tot}}(w^i)$
    $\phi^*(w) \leftarrow \frac{1}{m}\sum_{i=1}^m [k(w^i, w)\nabla_{w^i} \log p(w^i|\mathcal{O}) + \nabla_{w^i} k(w^i, w)]$
    $w^i \leftarrow w^i + \epsilon\phi^*(w^i)$

2) Likelihood Model for Modulation: Trajectory modulation is a crucial characteristic of movement primitives in novel situations, as the robot is expected to navigate through new desired points y_{tar}, commonly referred to as via-points, at specific times t. In various tasks, the ability to adapt to new starting or ending positions is necessary. Accordingly, we can apply Bayes' theorem in our SMP algorithms, similar to other movement primitive methods. In our approach, the conditioning distribution can be applied at any time point t in the trajectories by incorporating the desired observation for both single-modal and multi-modal trajectories, resulting in $p(w|y_{\text{tar}}) \propto p(y_{\text{tar}}|w)p(w)$. For single-modal trajectories, we express the likelihood term using an exponential utility function based on the optimality binary as a score function rather than a Gaussian distribution as done in [2]. To achieve this, we form a new via-point y_{tar} as an attractor at a specific time t and compute the difference between it and the coordinate of our reproduced trajectory as $F_t = y_t - y_{\text{tar}}$. Here, $y_t = \Phi_t w$ represents the D-dimensional vector of coordinates at time t in a trajectory reproduced by SMPs. Similarly to collision avoidance, we define the modulation cost as the inner product $\mathcal{C}_{\text{mod}} = \|F_t\|^2$. Thus, the likelihood of modulating a state to the target can be modelled as $\exp(-\lambda_{\text{mod}}\mathcal{C}_{\text{mod}})$ for $p(y_{\text{tar}}|w)$. If we want the trajectories to pass through a new via-point y_{tar} at a specific time step and also modulate to a new ending point y_{end}, the total likelihood term can be summed as $\exp(-\lambda_{\text{mod}}\mathcal{C}_{\text{mod}}) = \exp(-\lambda_t\|F_t\|^2 - \lambda_{\text{end}}\|F_{\text{end}}\|^2)$. This approach utilizes the utility function instead of computing the two KL-divergence objective functions separately, as done in [17]. Compared with modulation for single-modal trajectories, a GMM model is selected for the probabilistic likelihood term to modulate multi-modal trajectories. Given multiple via-points for modulation, we can model the H via-points using a GMM model as $p(y_{\text{tar}}|\theta_{\text{tar}}) = \sum_{h=1}^H \pi_h \mathcal{N}(y_{\text{tar}}|\mu_h, \Sigma_h)$ with parameters $\theta_{\text{tar}} = \{\mu_h, \Sigma_h\}_{h=1}^H$. We can drive the particle w to fit this GMM model, forming the cost function $\mathcal{C}_{\text{mod}} = \log p(\Phi_t w|\theta_{\text{tar}}) = \sum_{h=1}^H \pi_h \mathcal{N}(\Phi_t w|\mu_h, \Sigma_h)$. We can also sum different modulation costs together and express them as $\exp(-\lambda_{\text{mod}}\mathcal{C}_{\text{mod}}) = \exp(\lambda_t \log p(\Phi_t w|\theta_{\text{tar}}) - \lambda_{\text{end}}\|F_{\text{end}}\|^2)$. When we have multiple new goals to modulate for multi-modal trajectories, a similar method can be applied using the GMM of y_{end}. The Bayesian posterior for modulation can now be expressed as:

$$p_{\text{mod}}(w|\mathcal{O}) \propto \exp(-\lambda_{\text{mod}}\mathcal{C}_{\text{mod}})p(w). \tag{13}$$

3) SMPs with Smooth Box Prior: In addition to integrating the penalty term into the likelihood cost, we introduce a supplementary prior called the smoothed box prior, which functions as a smooth regularization or soft constraint. In our approach, we prioritize the empirical distribution over deterministic trajectories or parametric distributions of trajectories. Consequently, the particles must replicate the trajectories while adhering to constraints on the ending or starting points. For instance, with each iteration of SVGD, the trajectory's endpoint is likely to be constrained within a specified region around the target point. To facilitate this, we employ a smoothed approximation of a uniform prior, denoted as $p_{\text{box}}(w)$, which is represented as a box $B = \{\Phi_{\text{tar}}w : a \leq \Phi_{\text{tar}}w \leq b\}$, where $\Phi_{\text{tar}}w$ denotes the probability density function of this prior.

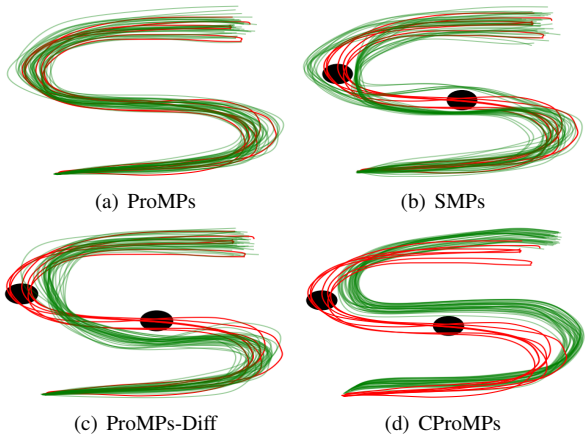
$$p_{\text{box}}(w) \propto \exp\left(-\frac{d(\Phi_{\text{tar}}w, B)^2}{\sqrt{2\sigma^2}}\right) \tag{14}$$

The distance function $d(\Phi_{\text{tar}}w, B)^2$ is defined as the minimum of $|\Phi_{\text{tar}}w - x'|$, where $x' \in B$. It fully supports the real numbers and is differentiable everywhere. If the endpoint does not lie within the region defined by the box B, the gradient flow of this prior can be sufficiently large to drive the trajectory points toward the box region. Once the trajectory points reach this area, the gradient flow becomes zero, similar to the behaviour defined by a uniform distribution. We can regard it as a variant of a barrier function but with both upper and lower bounds. This prior can be referred to as a hyper-prior and added to the existing $\log p(w)$ in the form $\log p_{\text{tot}}(w) = \log p(w) + \log p_{\text{box}}(\Phi_{\text{tar}}w)$. This smooth box prior can be applied during SVGD iterations, thereby constraining the coordinates within the required region. The results of including this smooth box prior are listed in Table I. We record the percentage of the endpoints from 42 samples of trajectories that lie within the required region when the smooth box prior is implemented.

TABLE I

THE PERCENTAGE OF TRAJECTORY ENDING POINTS THAT REMAIN WITHIN THE REQUIRED REGION UNDER THE SPECIFIED THRESHOLD

	With Smooth Box Prior	Without Smooth Box Prior
LASA Dataset	Within the region	Within the region
S Shape	42/42	40/42
S + W Shape	42/42	36/42



(a) ProMPs (b) SMPs (c) ProMPs-Diff (d) CProMPs

Fig. 1. The plots depict the alphabet character 'S' from the LASA dataset reproduced by the following methods: (a) ProMPs, (b) SMPs, (c) ProMPs-Diff, and (d) CProMPs. Demonstrations are highlighted in red, while reproduced trajectories generated by movement primitive methods are shown in green. Black solid circles represent obstacles.

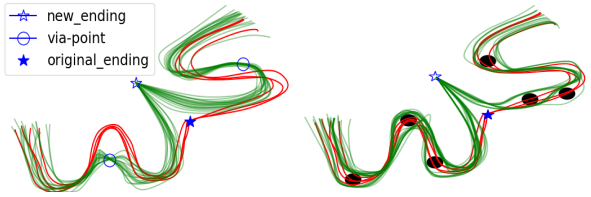
4) SVGD update for adaptation with Smooth Box Prior: Hence, we reformulate the KL optimization problem into an unconstrained optimization, amenable to a solution via SVGD for the tasks. Consequently, the log-posterior gradient for each particle can be written as the summation of the log-likelihood gradient and the log-prior gradient.

$$\nabla_{w^i} \log p(w^i|\mathcal{O}) = \nabla_{w^i}(-\lambda_{\text{obs}}\mathcal{C}_{\text{obs}} - \lambda_{\text{mod}}\mathcal{C}_{\text{mod}}) + \nabla_{w^i} \log p_{tot}(w^i) \tag{15}$$

With the implementation of SVGD, we can rewrite Eq.(11) :

$$\phi^*(w^i) = \frac{1}{m}\sum_{i=1}^{m}[k(w^i, w)(\nabla_{w^i}(-\lambda\mathcal{C}_{\text{obs}} - \lambda\mathcal{C}_{\text{mod}}) + \nabla_{w^i} \log p_{\text{tot}}(w^i)) + \nabla_{w^i} k(w^i, w)], \tag{16}$$

and the stein particles w^i can be updated by gradient with $w^i = w^i + \epsilon\phi^*(w^i)$.



(a) Modulation for S+W multi-modal dataset (b) Modulation and Collision Avoidance for S+W multi-modal dataset

Fig. 2. The plots display: (a) the modulation of a multi-modal case from the LASA dataset. Demonstrations are depicted in red, while reproduced results are shown in green. In both cases, the trajectories generated by the adapted primitives method are modulated to pass through via-points and reach a new ending point using our method; (b) the reproduced trajectories can be adapted to a new end point (red circle) from the old ending-point (blue circle) while avoiding obstacles (black solid circles) and successfully retaining the characteristics of the original multi-modal demonstrations. The demonstrations are shown in red, and the adapted trajectories are in green.

V. Experimental Evaluations

In this section, we evaluate our method with the LASA 2D dataset, robotics imitation learning in PyBullet simulation, and 6-DOF JACO manipulator in the real world.

Baseline : In Sec.V-A and Sec.V-B, SMPs are evaluated against the following three methods of adapting motion in novel environments.

- **CProMPS** [15]: In each iteration, following the reparameterization trick in GVI, we sample trajectories from the updated Gaussian distribution and subsequently calculate collision costs based on occupancy mapping.
- **ProMPT-O**: It is a ProMPs-based trajectory optimization method [14]. We implement it here as a baseline to test the adaptation of the reproduced trajectories.
- **ProMPs-Diff**: We follow a similar method as described in [7], integrating ProMPs with diffeomorphic transforms. We create a generalized imitation learning framework by adapting the mean of the Gaussian distribution of weights in ProMPs.

A. Experiment 1. LASA 2D dataset

1) Collision Avoidance for single modal demonstration: Firstly, we tested our methods and baseline on the 2D S-shape with the LASA Dataset, using 7 demonstrations and two obstacles. We visualize the qualitative results of the generated obstacle avoidance problems in Fig. 1. For each method, we created 42 particles. It is evident that our method can generate smooth multi-modal trajectories instead of being confined to a single-modal distribution, as observed in other methods. Additionally, our approach ensures that the reproduced trajectories have the highest similarity to the original demonstrations while maintaining an almost collision-free behaviour. This advantage stems from the fact that other adapted primitives are modelled into a single Gaussian model to obtain a closed-form solution, thereby restricting motion to navigate around obstacles within a single modal distribution and eventually deviating from the original movement primitives. In contrast, our method does not have these restrictions. It can present the adapted trajectories in multi-modal form by considering the information from the mapping of the new environment with the two existing obstacles, which can induce different homotopy classes.

2) Collision Avoidance and Modulation for multi-modal demonstration: After that, the qualitative results of the adaptation of multi-modal (combining S and W shapes) trajectories with the smooth box prior are presented in Fig. 2. We show that multi-modal trajectories can be modulated to new via-point(s) and a new ending point by SMPs with the smooth box prior in Fig. 2(a). In Fig. 2(b), all reproduced multi-modal trajectories generated by particles can successfully avoid the obstacles, and they can be modulated from the original final point (blue circle) to the new final point (red circle).

B. Experiment II: Robot Simulation

1) Collision Avoidance for Robot Manipulator: Firstly, we describe the tasks, metrics and baselines for our experi-

TABLE II

NUMERICAL EVALUATION OF COMPARING GENERATED TRAJECTORIES ACROSS 4 SCENARIOS, REPEATED 20 TIMES (MEAN ± STANDARD DEVIATION)

Method	Scenario	W.D	F.D	M.M.D	Succ.Rate	Avg.Traj.Len	Min.Traj.Len
Ours	Draw-S	**0.48%±0.05%**	**0.77±0.37**	**0.21±0.04%**	**98.61%±0.18%**	1.44±2.81%	1.34±4.36%
CProMPs[15]		0.75%±0.02%	0.82±0.11	0.51±0.01%	98.35% ±0.06%	**1.38±2.62%**	**1.22±2.18%**
ProMPT-O[14]		1.21%±0.13%	0.88±0.31	1.91±0.02%	95.27%±0.11%	1.40±3.74%	1.25±3.08%
ProMPs-Diff[7]		0.91%±0.01%	0.78±0.18	1.04±0.01%	91.5%±0.09%	1.39±1.52%	1.27±5.45%
Ours	Placement	1.91%±0.21%	**2.31±0.15**	**0.22±3.57%**	99.91%±0.07%	**1.39±1.80%**	**1.13±3.77%**
CProMPs		7.68%±0.34%	4.11± 0.17	0.61±3.83%	**100%±0.00%**	1.56±4.14%	1.24±2.03%
ProMPT-O		5.56%±0.34%	3.14±0.18	0.41±1.58%	99.82%±0.06%	1.61±1.14%	1.33±1.30%
ProMPs-Diff		**1.81%±0.27%**	2.38±0.07	**0.22±1.03%**	99.23%±0.11%	1.42±1.52%	1.24±3.22%
Ours	Push-Lift-Box	**3.34%±0.50%**	**2.45±0.14**	**0.10±2.00%**	99.34%±0.18%	**2.51±6.31%**	**2.31±3.05%**
CProMPs		5.80% ±0.43%	2.92±0.15	0.19±1.83%	**100%±0.00%**	2.61±1.98%	2.55±2.27%
ProMPT-O		10.63%±1.07%	4.11±0.27	0.51±1.71%	**100%±0.00%**	3.19±5.08%	3.25±8.23%
ProMPs-Diff		8.31%±0.64%	4.41±0.33	0.59±2.90%	**100%±0.00%**	3.90±2.15%	3.51 ±6.24%
Ours	Placement on Double-Sides	**1.56%±0.71%**	**2.04±0.07**	**0.14 ±2.05%**	**99.50%±0.12%**	**1.13±3.47%**	**0.95±2.62%**
CProMPs		16.12%±1.21%	12.25±1.31	1.47±1.60%	0.00%±0.00%	1.69±2.18%	2.16±4.57%
ProMPT-O		17.51%±3.02%	10.76±0.13	1.32±1.01%	0.00%±0.00%	2.21±4.19%	1.75±3.10%
ProMPs-Diff		15.41%±2.06%	11.81±0.09	1.21±0.77%	0.00%±0.00%	2.10±2.16%	1.55±2.48%

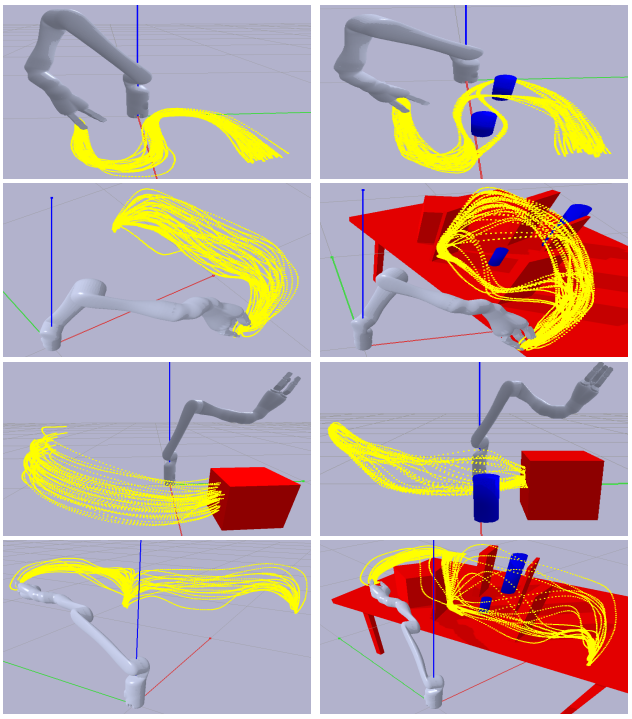
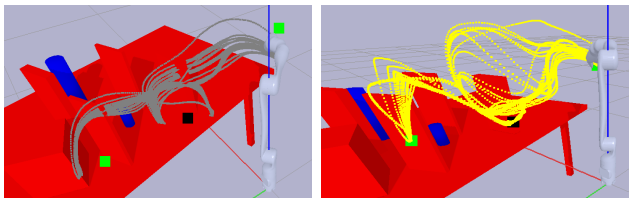


Fig. 3. Sampled trajectories of: (Left) original motion primitives; (Right) adapted by SMPs. Scenarios from top to bottom: S shape drawing; placement; push-lift box; placement on double sides (multi-modal). The robot end-effector stops at the desired ending point of trajectories.

ment.

Task : Then we implement the methods on a JACO robot in simulation to complete the following tasks: (i) The robot should draw an S-shape in 3D, starting from various nearby points and ending at a common point (7 demonstrations from the LASA Dataset extended to 3D). (ii) The robot should place an object in a motion planning benchmark scenario with a common starting and ending point (7 demonstrations in joint space collected from a real robot). (iii) The robot should start from different points within a region to push and lift a box to a final target position (7 demonstrations in joint space collected from a real robot). (iv) The robot should start from the middle and move to both sides to complete placement tasks in a motion bench marker scenario [23](14 demonstrations in joint space collected from a real robot). In Fig. 3, we illustrate the demonstration and reproduced robot motion with SMPs both in the absence of obstacles and in new scenarios with obstacles in Pybullet simulation.

Metric : There are six metrics used to compare the quality of reproduced trajectories :(i) Maximum Mean Discrepancy (M.M.D), a method to measure the discrepancy between the adapted trajectory distribution and initial trajectory distribution generated from original movement primitives. (ii) Frechet Distance (F.D) is a metric used to assess the similarity between two sets of trajectories generated by the adapted movement primitives and the original movement primitives separately. (iii) Wasserstein Distance (W.D), another method to measure the discrepancy between the two trajectory distributions or known as optimal transport distance (iv) Successful Rate (Succ.Rate), we compute how many adapted trajectories in the empirical distribution succeed in reaching the final goals and completing the tasks without collision. (vi) Average Trajectory Length (Avg.Traj.Len), we explore how the multi-modality can avoid local minimum and find better results. Thus, we investigate the trajectory length of adapted trajectories in meters. (vi) Min Trajectory Length



(a) Original Trajectories (b) Adapted Trajectories

Fig. 4. We Demonstrate that the multi-modal trajectories are modulated to new ending points on both sides (green square) while also beginning from a new common starting point (black square) by SMPs. (a) The multi-modal demonstrations of trajectories (Grey) for robots start with different starting points in the middle and are placed at two sides in a new scenario with many collisions. (b) Adapted collision-free multi-modal trajectories (yellow) generated by SMPs start from the common new starting point and placed at the new destinations at two sides of the desk

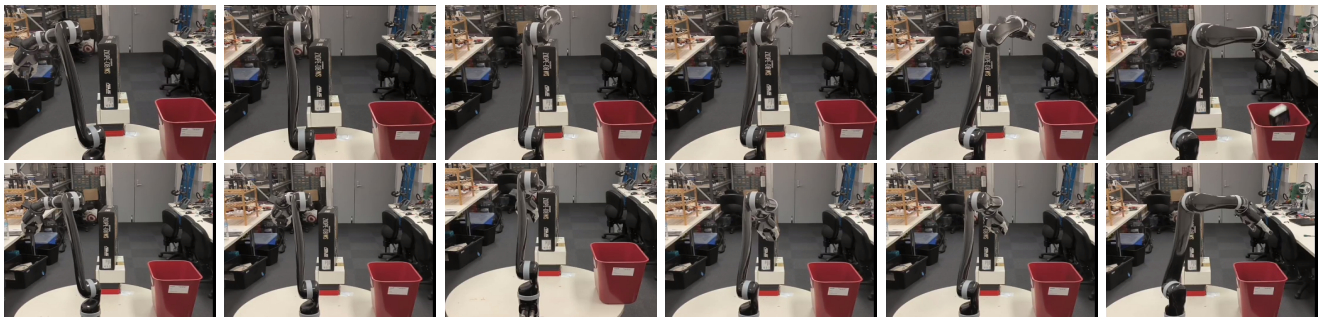


Fig. 5. (Top) One trajectory is generated by SMPs to navigate over the top of the obstacle, ensuring collision avoidance and subsequently placing the box into the trash bin. (Bottom) Conversely, another trajectory is produced to accomplish the same task, this time circumventing obstacles from the side.

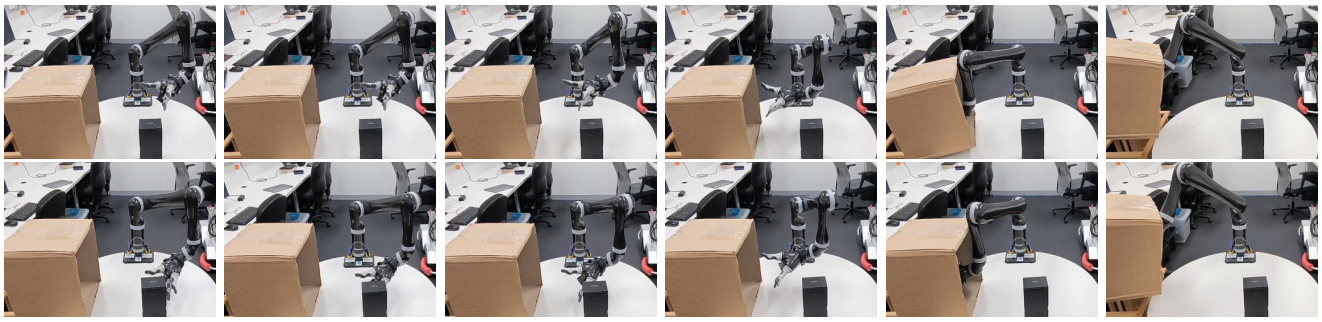


Fig. 6. (Top) The adapted robot motion ascends from the tops of obstacles to push and lift the box. (Bottom) The adapted robot motion maneuvers alongside obstacles to push and lift the box

(Min.Traj.Len), because our method can present multi-modal solutions, we check the minimum trajectory length to see how multi-modal can benefit.

Each demonstration is performed with a step size of 200. We select Gaussian basis functions with $K = 15$ for Φ. The quantitative results are presented in Table II. We conducted 20 tests, each consisting of 42 trajectory samples for all methods, and recorded the mean values. The variance in the results is small but is still included in the table. In the first three tasks, which involve single-modality demonstrations, we utilized a Gaussian distribution as the prior, derived from ProMPs [1], for both our method and the baseline. Without the Gaussian prior, the other baseline methods could not be implemented. From the results, across the two single-modality tasks ('Draw-S' and 'Push-Lift-Box'), our method consistently demonstrated the lowest values in W.D, F.D, and M.M.D compared to the other methods while maintaining a high success rate. This indicates that our method can achieve the highest similarity after adaptation while successfully completing tasks. In the 'Placement' task, SMPs yield results for the W.D and M.M.D metrics that are nearly comparable to those of ProMPs-Diff but with a higher success rate. In all three tasks, our method demonstrates the ability to represent different homotopy classes of trajectories caused by obstacles, as shown in Fig. 3. This capability arises from our method's design, which prioritizes similarity by incorporating environmental information and presenting trajectories in a multi-modal distribution. In contrast, other baseline methods focus solely on avoiding obstacles within a single-modal distribution. Our method's multi-modal property enables it to discover superior solutions, such as achieving the lowest average trajectory length by effectively avoiding local minima, except in the 'Draw-S' task. In this particular task, other methods attempt to evade obstacles within a single Gaussian model, making it easier for them to deviate from the original movement primitives, leading to lower similarity. However, this can also result in reduced trajectory lengths in this specific task. For the multi-modal task 'Placement on Double Sides,' where demonstrations are multi-modal, we select an MM with two components as the prior for SMPs since our method is not restricted to using a Gaussian distribution for the prior or candidate distribution. Other baselines cannot be used for adapting multi-modal demonstrations with a GMM prior [18]. Based on the findings presented in Table II, it is evident that SMPs consistently achieve high similarity with the original movement primitives, with an almost 100% success rate. The results obtained by SMPs for the three metrics—W.D, F.D, and M.M.D—are still reasonable. Conversely, other methods fail due to their incapacity to address challenges presented in multi-modal scenarios. The primary advantage of our method lies in its ability to adapt multi-modal trajectories.

2) Collision Avoidance with Modulation for Robot Manipulator: In this section, we aim to assess the ability of our methods to modulate motion in a multi-modal distribution towards multiple new ending points and a new common starting point while avoiding obstacles. We conducted our analysis using the JACO Robot for the motion planning benchmark. This time, we introduce a shift in the demonstrations, resulting in different starting points. Despite this variation,

we maintain consistency in our approach by employing the same settings for movement primitives and priors as used in the collision avoidance task. We set a new common starting point for the two-modal trajectories. Additionally, the ending point on the left side is shifted towards the middle of the desk by 15 cm, while the ending point on the right side is lifted by 15 cm. We highlight the two ending points as green squares and the new common starting point as a black square. In Fig. 4, we present that our method can modulate the two modal primitives to the new ending points. Even when the demonstrations start at different positions initially, they can still be modulated to the new common starting point, especially with the implementation of the smooth box prior. Additionally, our method can adapt multi-modal trajectories to avoid obstacles simultaneously while maintaining similarity to the original distributions.

C. Experiment III: Real Robot Experiment

To evaluate our method with a real robot, we conducted two tasks: 1. 'Placement' and 2. 'Pushing-Lifting-Box' using a JACO arm. In each task, the robot arm moves to avoid obstacles, seamlessly executing multi-modal adapted trajectories generated from updated primitive without any collisions, ultimately accomplishing the tasks successfully. We present our tests in Fig. 5, Fig. 6. In the first task, we showed that the robot could move in two models to avoid obstacles and place the box into the bin. In the second task, we also presented the robot that can push and lift the box while avoiding the obstacle in two trajectory models.

VI. Conclusion

In this paper, we presented Stein Movement Primitives, a framework that utilizes Stein Variational Gradient Descent to estimate an empirical, non-parametric distribution of adapted trajectories. Our method enables both single-modal and multi-modal robot movement primitives to adapt to new scenarios seamlessly. Moreover, SMPs have the capability to generate multi-modal trajectories with different homotopy classes by including information on obstacles in the likelihood term. It can be used for navigation in complex environments while avoiding local minima by reasoning about multiple trajectories at the same time, a significant improvement over previous primitive movement methods. Additionally, the ability of SMPs to modulate to new target points in multi-modal trajectories is another unique feature. Experimental validation conducted in simulation and real-world manipulation tasks demonstrates the superior performance of our approach compared to previous work. In the future, we plan to extend the method to address trajectory planning in dynamic environments. Our method is computationally inexpensive and fast when executed in the GPU and can potentially be applied to real-time dynamical scenarios in future work.

References

[1] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, 'Dynamical movement primitives: Learning attractor models for motor behaviors,' *Neural Computation*, vol. 25, no. 2, pp. 328–373, 2012.

[2] A. Paraschos, C. Daniel, J. R. Peters, and G. Neumann, 'Probabilistic movement primitives,' in *Proc. Int. Conf. Neural Inf. Process. Syst*, 2013, pp. 2616–2624.

[3] H. Hoffmann, P. Pastor, D. Park, and S. Schaal, 'Biologically-inspired dynamical systems for movement generation,' in *Proc. IEEE Int. Conf. Robot. Autom*, 2009, pp. 2587–2592.

[4] R. Krug and D. Dimitrov, 'Model predictive motion control based on gen eralized dynamical movement primitives,' *J. Intell. Robot. Syst.*, vol. 77, no. 1, pp. 17–35, 2015.

[5] M. Ginesi, D. Meli, A. Calanca, D. Dall'Alba, N. Sansonetto, and P. Fiorini, 'Dynamic movement primitives: Volumetric obstacle avoidance,' in *Proc. 19th Int. Conf. Adv. Robot.*, 2009, pp. 234–239.

[6] A. Duan, R. Camoriano, D. Ferigo, D. Calandriello, L. Rosasco, and D. Pucci, 'Constrained dmps for feasible skill learning on humanoid robots,' in *Proc. IEEE-RAS Int. Conf. Humanoid Robots.*, 2018, pp. 1–6.

[7] W. Zhi, T. Lai, L. Ott, and F. Ramos, 'Diffeomorphic transforms for generalised imitation learning,' in *Proc. Learning for Dynamics and Control Conference*, 2022.

[8] A. Paraschos, R. Lioutikov, J. Peters, and G. Neumann, 'Probabilistic prioritization of movement primitives,' *IEEE Robot. Automation. Letter*, vol. 2, no. 4, pp. 2294–2301, 2017.

[9] S. Calinon, 'A tutorial on task-parameterized movement learning and retrieva,' vol. 9, no. 1, pp. 1–29, 2016.

[10] J. Silvério, S. Calinon, L. Rozo, and D. G. Cal, 'Learning task priorities from demonstration,' vol. 42, no. 3, pp. 78–94, 2019.

[11] D. Koert, G. Maeda, R. Lioutikov, G. Neumann, and J.Peters, 'Demonstration based trajectory optimization for generalizable robot,' in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2016, pp. 515–522.

[12] A. Colomé and C. Torr, 'Demonstration-free contextualized probabilis- tic movement primitives, further enhanced with obstacle avoidance,' in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 3190–3195.

[13] D. Koert, J. Pajarinen, A. Schotschneider, S. Trick, C. Rothkopf, and J. Peter, 'Learning intention aware online adaptation of movement primi- tives,' in *IEEE Robot. Autom. Lett*, 2019, pp. 3719–3726.

[14] T. Loew, T. Bandyopadhyay, J. Williams, and P. Borges, 'Prompt: Probabilistic motion primitives based trajectory planning,' in *Robotics Science and Systems (RSS)*, 2021.

[15] F. Frank, A. Paraschos, P. Smagt, and B. Cseke, 'Constrained probabilistic movement primitives for robot trajectory adaptation,' *IEEE Trans. Rob*, vol. 38, no. 4, pp. 2276–2294, 2022.

[16] E. Pignat and S. Calinon, 'Bayesian gaussian mixture model for robotic policy imitation,' *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4452–4458, 2019.

[17] Y. Huang, L. Rozo, J. Silvério, and D. G. Caldwell, 'Kernelized movement primitives,' *IJRR*, vol. 38, no. 7, pp. 297–304, 2019.

[18] W.Zhi, L.Ott, and F. Ramos, 'Probabilistic trajectory prediction with structural constraints,' in *Proc. of IROS*, 2021.

[19] Q. Liu and D. Wang, 'Stein variational gradient descent: A general purpose bayesian inference algorithm,' in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 2378–2386.

[20] A. Lambert, F. Ramos, B. Boots, D. Fox, and A. Fishman, 'Stein variational model predictive control,' in *Proc. of CoRL*, 2020.

[21] L. Barcelos, A. Lambert, R. Oliveira, P. Borges, B. Boots, and F. Ramos, 'Dual online stein variational inference for control and dynamics,' in *Proc. of RSS*, 2021.

[22] A. Lambert, B. Hou, R. Scalise, S. Srinivasa, and B. Boots, 'Stein variational probabilistic roadmaps,' in *Proc. IEEE Int. Conf. Robot. Autom (ICRA).*, 2023.

[23] C. Chamzas, C. Quintero-Peña, Z. Kingston, A. Orthey, D. Rakita, M. Gleicher, M. Toussaint, and L. E. Kavraki, 'Motionbenchmaker: A tool to generate and benchmark motion planning datasets,' *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 882–889, 2022.