

Memory of Motion for Warm-starting Trajectory Optimization

Teguh Santoso Lembono^{1,2}, Antonio Paolillo¹, Emmanuel Pignat^{1,2},
and Sylvain Calinon^{1*}

July 3, 2019

Abstract

Trajectory optimization for motion planning requires a good initial guess to obtain good performance. In our proposed approach, we build a memory of motion based on a database of robot paths to provide a good initial guess online. The memory of motion relies on function approximators and dimensionality reduction techniques to learn the mapping between the task and the robot paths. Three function approximators are compared: k -Nearest Neighbor, Gaussian Process Regression, and Bayesian Gaussian Mixture Regression. In addition, we show that the usage of the memory of motion can be improved by using an ensemble method, and that the memory can also be used as a metric to choose between several possible goals. We demonstrate the proposed approach with the motion planning on a dual-arm PR2 robot.

1 Introduction

Motion planning for robot with high Degree-of-Freedoms (DoFs) presents many challenges, especially in the presence of constraints such as obstacle avoidance, joint limits, etc. To handle the high-dimensionality and the various constraints, many works focus on *trajectory optimization* methods that attempt to find a locally optimal solution. In this approach, the motion planning problem is formulated as an optimization problem

$$\min_{\mathbf{q}_{0:T}} \ell(\mathbf{q}_{0:T}), \quad \text{s.t.} \quad \mathbf{g}(\mathbf{q}_{0:T}) \leq \mathbf{0}, \quad \mathbf{h}(\mathbf{q}_{0:T}) = \mathbf{0}, \quad (1)$$

where $\mathbf{q}_{0:T}$ denotes the robot's configurations from time step $t = 0$ to $t = T$; $\ell(\cdot)$, $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$ are the cost, the inequality and the equality constraints. The solution of (1) is the *path* $\mathbf{y} = \mathbf{q}_{0:T}^* = (\mathbf{q}_0^{*\top}, \dots, \mathbf{q}_T^{*\top})^\top \in \mathbb{R}^{DT}$, with D the dimension of \mathbf{q}_t . When the path is parameterized by time, it is called *trajectory*.

As an example, consider the planning problem depicted in Fig. 1, where the PR2 robot has to move its base around an object or to perform a dual-arm motion to pick items from the shelves. If the task $\mathbf{x} = (\mathbf{q}_{\text{init}}^\top, \mathbf{q}_{\text{goal}}^\top)^\top$ is to move from an initial configuration \mathbf{q}_{init} to a goal configuration \mathbf{q}_{goal} while minimizing the total joint velocity, the optimization problem can be written as

$$\min_{\mathbf{q}_{0:T}} \sum_{t=0}^{T-1} \|\mathbf{q}_{t+1} - \mathbf{q}_t\|^2, \quad \text{s.t.} \quad \mathbf{q}_0 = \mathbf{q}_{\text{init}}, \quad \mathbf{q}_T = \mathbf{q}_{\text{goal}}. \quad (2)$$

Other constraints can also be added, e.g. to avoid collisions, to comply with joint limits, etc.

Such optimization problems are in general non-convex, especially due to the collision constraints, which makes finding the global optimum very difficult. Trajectory optimization methods such as TrajOpt [1], AICO [2], CHOMP [3], or STOMP [4] solve the non-convex problem by iteratively optimizing around the current solution. While such approach is very popular and yields good practical results, the convergence and the quality of the solution are very sensitive to the choice of the initial guess. If it is far from the optimal solution, the method can get stuck at a poor local optimum.

^{*1} Idiap Research Institute, Martigny, Switzerland ² EPFL, Lausanne, Switzerland

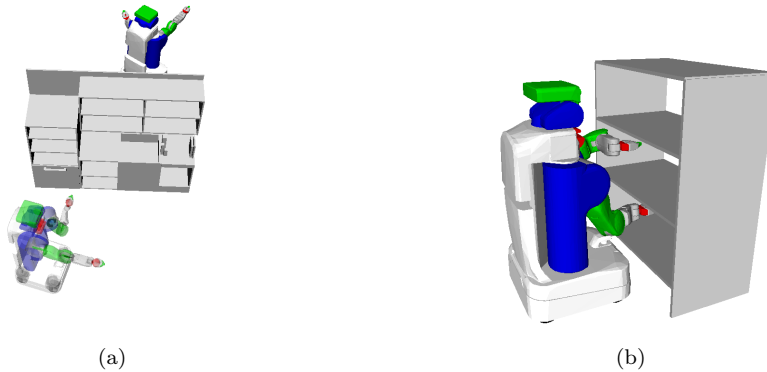


Figure 1: Examples of motion planning problem with the PR2 robot: (a) moving the base between two points while avoiding an obstacle and (b) dual arm motion to pick items from a shelf to another.

To overcome this problem, our approach builds a *memory of motion* that learns how to provide a good initialization (i.e., a *warm-start*) to the solver based on previously solved problems. Functionally, the memory of motion is expected to learn the mapping $\mathbf{f} : \mathbf{x} \rightarrow \mathbf{y}$ that maps each task \mathbf{x} to the robot path \mathbf{y} in the database. Such mapping can be highly nonlinear and *multimodal* (i.e., one task \mathbf{x} can be associated to several robot paths \mathbf{y}), and the dimension of \mathbf{y} is typically very high. Our proposed method relies on machine learning techniques such as function approximation and dimensionality reduction to learn this mapping effectively. Note that we use the term *memory of motion* to denote not only the database of motions but also the algorithms used to predict the warm-start.

The contribution of this paper is the following. First, we propose the use of function approximation methods to learn the mapping $\mathbf{f}(\mathbf{x})$. We consider three methods: k -Nearest Neighbor (k -NN), Gaussian Process Regressor (GPR) and Bayesian Gaussian Mixture Regression (BGMR), and discuss the different characteristics on various planning problems. We show in particular that BGMR handles multimodal output very well. Furthermore, we demonstrate that using an ensemble of function approximators to provide the warm-start outperforms any individual function approximator. Finally, we show that the resulting memory of motion can be used as a metric to choose optimally from several possible goals.

The paper is organized as follows. In Section 2 we discuss the related work that use the concept of memory of motion for various problems. Section 3 explains the methods for constructing and using the memory of motion. The experimental results are presented and discussed in Section 4. Finally, Section 5 concludes the paper.

2 Related Work

The idea of using a memory of motion to warm-start an optimization solver has previously been explored in the context of optimal control and motion planning. In [5] a trajectory library is constructed to learn a control policy. A set of trajectories are planned offline using A^* algorithm and stored as library, then Nearest-Neighbor is used online to determine the action to perform at each state. In [6], they use similar approach to predict an initial guess for balancing a two-link robot, which is then optimized by Differential Dynamic Programming [7]. An iterative method to build a memory of motion to initialize an optimal control solver is proposed in [8]. They use neural networks to approximate the mapping from the task descriptors (initial and goal states) to the state and control trajectories. Another neural network is trained to approximate the value function, which is then used as a metric to determine how close two states are dynamically. In [9] Gaussian Process Regression (GPR) is used to predict new trajectories based on the library of demonstrated movements encoded as Dynamic Movement Primitives (DMP) [10]. GPR is used to map the task descriptors to the DMP parameters.

In robot motion planning, Probabilistic Roadmap (PRM) [11] can be seen as the construction of

memory of motion by precomputing the graph connecting robot configurations. Some works exploit Rapidly-exploring Random Trees (RRT) [12], another popular sampling-based method. For example, in [13] an offline computation is used to speed-up the online planning in the form of an additional bias in sampling the configurations. The *Lightning* framework is proposed in [14] to plan paths in high-dimensional spaces by learning from experience. The path library is constructed incrementally. Given the current path library and a task to be executed, the algorithm runs two versions of the planner online, one that plans from scratch and the other one initialized by the library. In [15] a high-dimensional (791) task descriptor is constructed, which consists of the initial and target robot configurations and the pairwise distance between the obstacles and the robot body parts. Given a current task descriptor, they use k -Nearest Neighbor and Locally Weighted Regression to retrieve the initial trajectory, which is then optimized by a trajectory optimizer. The metric between the task descriptors is refined to minimize the necessary refinement of the initial trajectory using L_1 norm, such that they obtain a sparse metric and hence sparse descriptors. In [16], a subindexing is used to reduce the amount of memory storage and to use the subtrajectories of the original solutions. In addition, for a given task descriptor that is associated to a trajectory solution, the algorithm samples new task descriptors from the region near the current descriptor, and runs the trajectory optimization using the current trajectory solution as the initial guess. If the optimization converges to the good solution, the new task descriptor is mapped to the current trajectory solution.

As compared to the above works, our proposed method has the following differences: (i) none of the above methods attempt to handle the multimodal output cases. We show that BGMR can handle well such cases, (ii) we show that using an ensemble of methods to provide the warm-start outperforms the individual methods, and (iii) we show that using the memory to select between different goals can improve the solver performance significantly.

3 Method

Section 3.1 discusses the main idea of building the memory of motion using function approximation and dimensionality reduction techniques to learn the mapping between the task and the associated robot path. Section 3.2 then explains how the performance of the memory motion can be improved by using an ensemble method. Finally, Section 3.3 explains how the memory of motion can be used as a metric to choose between different goals.

3.1 Function Approximations and Dimensionality Reduction Techniques

To learn the mapping $\mathbf{f} : \mathbf{x} \rightarrow \mathbf{y}$, we firstly generate a large number of tasks \mathbf{x} and the corresponding robot paths \mathbf{y} . This is done by sampling \mathbf{x} from the space of possible tasks and run the trajectory optimizer to obtain the robot paths \mathbf{y} until we obtain N samples (\mathbf{x}, \mathbf{y}) . Let $\mathbf{X} = (\mathbf{x}_0, \dots, \mathbf{x}_{N-1})$ and $\mathbf{Y} = (\mathbf{y}_0, \dots, \mathbf{y}_{N-1})$. The mapping \mathbf{f} can then be learned by training function approximators using the database $\{\mathbf{X}, \mathbf{Y}\}$. In this paper we consider three function approximators: k -NN, GPR, and BGMR.

3.1.1 k -Nearest Neighbor (k -NN)

k -NN is a very simple non-parametric method. Given a task \mathbf{x}^* , the algorithm finds K samples $\{\mathbf{x}_k, \mathbf{y}_k\}_{k=1}^K$ in the database $\{\mathbf{X}, \mathbf{Y}\}$ where $\{\mathbf{x}_k\}_{k=1}^K$ are the K -nearest to \mathbf{x}^* according to a chosen metric (in this paper the Euclidean metric is used). It then predicts the corresponding robot path \mathbf{y}^* by taking the average $\mathbf{y}^* = \frac{1}{K} \sum_{k=1}^K \mathbf{y}_k$. The method is very simple to implement and it works well if there is a sufficiently dense dataset, but it suffers from the curse of dimensionality; as the dimension of \mathbf{x} increases, the number of data that needs to be stored increases exponentially. This method is mainly considered as the baseline against the next two methods.

3.1.2 Gaussian Process Regressor (GPR)

Like k -NN, GPR [17] is a non-parametric method which improves its accuracy as the number of data increases. While having higher computational complexity as compared to k -NN, GPR tends to better interpolate, resulting in higher approximation accuracy. Given the database $\{\mathbf{X}, \mathbf{Y}\}$, GPR assigns a Gaussian prior to the joint probability of \mathbf{Y} , i.e., $p(\mathbf{Y}|\mathbf{X}) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{X}), \mathbf{K}(\mathbf{X}, \mathbf{X}))$. $\boldsymbol{\mu}(\mathbf{X})$ is the mean function and $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is the covariance matrix constructed with elements $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, where $k(\mathbf{x}_i, \mathbf{x}_j)$ is the kernel function that measures the similarity between the inputs \mathbf{x}_i and \mathbf{x}_j . In this paper we use Radial Basis Function (RBF) as the kernel function, and the mean function $\boldsymbol{\mu}(\mathbf{X})$ is set to be zero as usually done in GPR.

To predict the output \mathbf{y}^* given a new input \mathbf{x}^* , GPR constructs the joint probability distribution of the training data and the prediction, and then conditions on the training data to obtain the predictive distribution of the output, $p(\mathbf{y}^*|\mathbf{x}^*) \sim \mathcal{N}(\mathbf{m}, \boldsymbol{\Sigma})$, where \mathbf{m} is the posterior mean computed as

$$\mathbf{m} = \mathbf{K}(\mathbf{x}^*, \mathbf{X}) \mathbf{K}^{-1}(\mathbf{X}, \mathbf{X}) \mathbf{Y}(\mathbf{X}), \quad (3)$$

and $\boldsymbol{\Sigma}$ is the posterior covariance which provides a measure of uncertainty on the output. In this work we simply use the posterior mean \mathbf{m} as the output, i.e., $\mathbf{y}^* = \mathbf{m}$.

While having good approximation accuracy, one major limitation with GPR is that it does not scale well with very large datasets. There are variants of GPR that attempt to overcome this problem, e.g., sparse GPR [18] or using Stochastic Variational Inference (SVI) [19]. More details on GPR can be found in [17] and [20].

3.1.3 Bayesian Gaussian Mixture Regression (BGMR)

When using RBF as the covariance function, GPR assumes that the mapping from \mathbf{x} to \mathbf{y} is smooth and continuous. When this assumption is met it performs very well, but otherwise it will yield poor results. For example, when there is discontinuity in the mapping or there are multimodal outputs, GPR tends to average the solutions from both sides of the discontinuity or from both modes. This characteristic is also shared by many other function approximators. To handle discontinuity and multimodality problems, using local models is one of the possible solutions. Each local model can be fit to each side of the discontinuity or to each mode.

Gaussian Mixture Regression (GMR) is an example of such local models approaches [21]. It can be seen as a probabilistic mixture of linear regressions. Given the database $\{\mathbf{X}, \mathbf{Y}\}$ it can be used to construct the joint probability of (\mathbf{x}, \mathbf{y}) as a mixture of Gaussians

$$p(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^K \pi_k \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (4)$$

where π_k , $\boldsymbol{\mu}_k$, and $\boldsymbol{\Sigma}_k$ are the k -th component's mixing coefficient, mean, and covariance, respectively. Given a query \mathbf{x}^* , the conditional probability of the output \mathbf{y}^* is also a mixture of Gaussians.

In GMR, the parameters π_k , $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are determined from the data by Expectation-Maximization method, while the number of Gaussians K is usually determined by the user. Bayesian GMR (BGMR) [22] is a Bayesian extension of GMR that allows us to estimate the posterior distribution of the mixture parameters (instead of relying on a single point estimate as in GMR). The number of components K can also be automatically determined from the data. As a Bayesian model, BGMR gives priors to the parameters π_k , $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$, and computes the posterior distribution of those parameters given the data. The prediction \mathbf{y}^* , given the input \mathbf{x}^* , is then computed by marginalizing over the posterior distribution and conditioning on \mathbf{x}^* . The resulting predictive distribution of \mathbf{y} is a mixture of t-distributions,

$$p(\mathbf{y}|\mathbf{x}^*, \mathbf{X}, \mathbf{Y}) = \sum_{k=1}^K p(k|\mathbf{x}^*, \mathbf{X}, \mathbf{Y}) p(\mathbf{y}|k, \mathbf{x}^*, \mathbf{X}, \mathbf{Y}), \quad (5)$$

where $p(k|\mathbf{x}^*, \mathbf{X}, \mathbf{Y})$ is the probability of \mathbf{x}^* belonging to the k -th component of the mixture, and $p(\mathbf{y}^*|k, \mathbf{x}^*, \mathbf{X}, \mathbf{Y})$ is a multivariate t-distribution, the mean of which is linear in \mathbf{x}^* . We can interpret (5) as K probabilistic linear regression models, each of which has the probability of $p(k|\mathbf{x}^*, \mathbf{X}, \mathbf{Y})$.

To obtain a point-prediction \mathbf{y}^* from (5), there are several approaches. One of the most used is to take the mean of the predictive distribution in (5) using moment matching. While this approach can provide smooth estimates (as required in many applications), the same problems as in GPR will appear in the case of discontinuity and multimodality; taking average in those cases will give us poor results. Instead, we propose to take, as the point prediction, the mean¹ of the component in (5) having the highest probability, which approximately corresponds to the mode of the multimodal distribution. Alternatively, we can also use the mean of each t-distributions as separate predictions, which gives us several possible solutions. In some cases (e.g., when we would like to retrieve all possible solutions) this approach can be very useful, as will be presented in Section 4.1.

3.1.4 Dimensionality reduction

In our problem, the path $\mathbf{y} \in \mathbb{R}^{DT}$ is a vector consisting of the sequence of configurations with dimension D during T time steps, which can be very high. This motivates us to use dimensionality reduction techniques to reduce the dimension of \mathbf{y} . For example, when T is large and the time interval is small, RBF can be used to represent the evolution of each degree of freedom as weights of the basis functions. Techniques such as Principal Component Analysis (PCA), Independent Component Analysis, Factor Analysis, and Variational Autoencoder [20] [23] can also be used. The mapping to be learned then becomes the mapping from \mathbf{x} to $\hat{\mathbf{y}}$, where $\hat{\mathbf{y}}$ is the projection of \mathbf{y} to the lower dimensional subspace. The advantage is that the memory required to store the data is reduced significantly, while the approximation performance is maintained or even improved because the important correlations between the variables are preserved. In this work, since the number of time steps is not large, we choose to use PCA to reduce the dimension of \mathbf{y} .

3.2 Using Ensemble Method to Provide Warm-Start

In machine learning, methods such as AdaBoost [24] and Random Forests [25] have shown that using an ensemble of methods often yields improved performances as compared to choosing a single method. We propose to use a simple ensemble method where we run multiple trajectory optimizations in parallel, each one warm-started by one of the function approximators in Section 3.1, and once one of them finds a successful path the others are terminated. Since each function approximator has different learning characteristics, combining them in this way can significantly improve the motion planning performance.

3.3 Using the Memory as a Metric

In some planning problems, there can be several alternative goals to be achieved. For example, in robot drilling task [26], the orientation around the drilling axis is free (the number of possible goals is infinite). A simple way is to choose one of the goals randomly, plan the motion, and if it fails then select another goal. While this is simple to implement, it does not make use of the benefit of having multiple goals. Another method is to plan the paths to each goal and select the one having the smallest cost, but this is computationally expensive. It will be useful, then, to have a metric that measures the cost to a given goal. Our proposed idea is to use the memory of motion as the metric.

In Section 3.1, the memory of motion learns how to predict an initial guess to achieve a task \mathbf{x} . The possible goals can then be formulated as multiple tasks $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{M-1}\}$. For each task \mathbf{x}_i , the memory of motion predicts the initial guess $\tilde{\mathbf{y}}_i$ corresponding to the task, and the cost $\ell(\tilde{\mathbf{y}}_i)$ can be computed. The initial guess $\tilde{\mathbf{y}}_i^*$ and the corresponding task \mathbf{x}_i^* with the lowest cost is then taken as the chosen goal to be given to the trajectory optimizer. If the computation of the cost can be done quickly relative to the time required for solving the optimization, such as the total discrete velocity in (2), this approach can yield significant improvements to the trajectory optimizer performance.

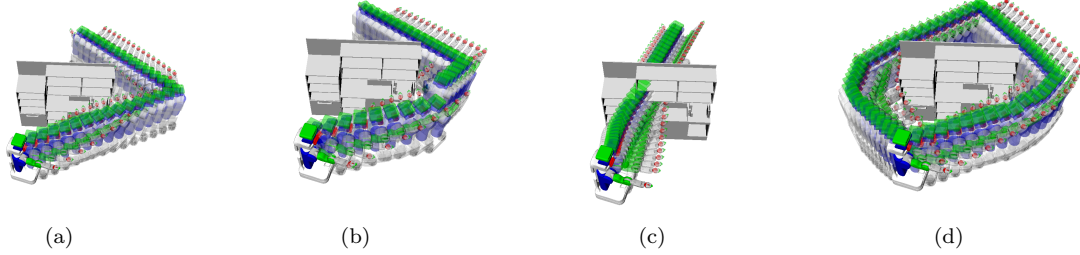
¹As in Gaussian distribution, the mean of a multivariate t-distribution is also its mode.

Table 1: Base motion planning, one waypoint.

Method	Success (%)	Conv. time (s)	Cost (rad/s)
STD	81.0	0.52	1.457
k -NN	97.0	0.32	1.480
GPR	95.0	0.32	1.388
BGMR	97.0	0.31	1.412

Table 2: Base motion planning, two waypoints.

Method	Success (%)	Conv. time (s)	Cost (rad/s)
STD	84.0	0.45	1.542
k -NN	96.0	0.35	1.666
GPR	00.0	-	-
BGMR	96.0	0.27	1.477

Figure 2: Motion planning for the PR2 mobile base. Warm-start produced by (a) straight-line with waypoint, (b) k -NN, (c) GPR and (d) BGMR.

4 Experiments

To evaluate the proposed method, we use several examples of motion planning for PR2 robot (see Fig. 1). TrajOpt [1] is used as the trajectory optimizer to be warm-started. The output is the robot path that accomplishes the given task. In this paper we only work with robot path as the output, but the method can be applied to robot trajectory as well.

We consider four motion planning cases presented in an ascending order of complexity. Each case is chosen to demonstrate certain characteristics of the proposed method. For each case, we follow the following procedures. First we generate the dataset by sampling N_{train} tasks and run TrajOpt to find the paths achieving the tasks. The number of time steps T is set to 30. In all cases, the cost is defined as the discrete velocity of the states, as defined in (2). The number of N_{train} is different for each case, depending on the complexity of the task. The function approximators are then trained with or without PCA using the dataset. We heuristically set 50 components for the PCA; for the k -NN, we use $K = 1$.

To validate the performance, we sample N_{test} random tasks and use the various methods to warm-start TrajOpt. The solutions are compared in terms of *convergence time*, *success rate* and *cost*. The planning is considered successful if the solution is feasible. In the presented results, we use the label ‘STD’ to refer to the solution obtained by warm-starting the solver with a straight-line path (via waypoint, if any), and the names of the function approximators for the rest. The subscript ‘PCA’ is added when the PCA is used. The query time for predicting the warm-starts by each method is negligible w.r.t. the convergence time (less than 50 ms), so they are not included in the comparison.

4.1 Base motion planning

The task is to plan the motion for the PR2 mobile base from a random pose in front of the kitchen to another random pose behind the kitchen (Fig. 1a). In this case, the state \mathbf{q} is the 3 DoF planar pose of the base. The task parameter is then $\mathbf{x} = (\mathbf{q}_{\text{init}}^\top, \mathbf{q}_{\text{goal}}^\top)^\top$. The database is constructed with $N_{\text{train}} = 200$ samples and the evaluation is performed with $N_{\text{test}} = 100$. Although this is an easy problem, TrajOpt actually finds it difficult to solve without a proper initialization. For example, initializing TrajOpt with a straight-line interpolation from \mathbf{q}_{init} to \mathbf{q}_{goal} never manages to find a feasible solution because it results in a path that moves the robot through the kitchen while colliding, and the solver get stuck in poor local optima due to the conflicting gradients. To obtain better initialization for building the database, we initialize TrajOpt with two manually chosen waypoints on the left and on the right of

Table 3: Planning from fixed \mathbf{q}_{init} to random \mathbf{q}_{goal} .

Method	Success (%)	Conv. time (s)	Cost (rad/s)
STD	80.0	0.76	1.831
k -NN	91.2	0.57	1.928
GPR	92.0	0.62	1.840
GPR _{PCA}	94.0	0.62	1.826
BGMR	90.8	0.62	1.854
BGMR _{PCA}	90.8	0.63	1.841

Table 4: Planning from random \mathbf{q}_{init} to \mathbf{q}_{goal} .

Method	Success (%)	Conv. time (s)	Cost (rad/s)
STD	71.6	0.81	1.211
k -NN	72.4	1.05	1.420
GPR	82.4	0.77	1.234
GPR _{PCA}	84.4	0.76	1.239
BGMR	76.0	0.77	1.274
BGMR _{PCA}	74.8	0.78	1.240
Ensemble	98.0	1.00	1.317

the kitchen (\mathbf{q}_{left} and $\mathbf{q}_{\text{right}}$, respectively).

We consider two cases of building the database: in the first one, we only use $\mathbf{q}_{\text{right}}$ as waypoint, while in the second we use both \mathbf{q}_{left} and $\mathbf{q}_{\text{right}}$. We initialize TrajOpt with the straight-line motion from \mathbf{q}_{init} to the waypoint and from the waypoint to \mathbf{q}_{goal} . With this setting we build the database, train the function approximators, and obtain the results shown in Table 1 and 2.

In the first case, the mapping from \mathbf{x} to \mathbf{y} is unimodal because all movements go through the right. Table 1 shows that the performance of k -NN, GPR and BGMR are quite similar. In the second case, however, the output is multimodal because the database contains two possible ways (modes) to accomplish the same task. This affects GPR significantly (see Table 2), as GPR takes the average of both modes and outputs a path that goes through the kitchen, while k -NN and BGMR are not affected. k -NN does not average the modes because we use $K = 1$, while BGMR overcomes the multimodality by constructing local models for each mode automatically.

Fig. 2 shows the examples of warm-starts produced by each method in the second case. As expected, GPR provides a warm-start that goes through the kitchen (hence 0 success rate). With BGMR, if we retrieve the components with the two highest probability, both possible solutions are obtained.

4.2 Planning from a fixed initial configuration to a random goal configuration

In this case, the state \mathbf{q} consists of 14 joint angles of the two 7 DoFs arms of PR2. The task is to move from a fixed \mathbf{q}_{init} to a random goal configuration \mathbf{q}_{goal} . The task \mathbf{x} can then be defined as the goal configuration ($\mathbf{x} = \mathbf{q}_{\text{goal}}$). The database is constructed with $N_{\text{train}} = 500$, and the evaluation results with $N_{\text{test}} = 250$ are presented in Table 3.

Since each PR2 arm is redundant, the path from \mathbf{q}_{init} to \mathbf{q}_{goal} can be multimodal, which may pose a problem for GPR. However, Table 3 shows that GPR and BGMR perform similarly. This is due to the fact that, although redundant robots can achieve a goal configuration in many different ways, planning using optimization results in similar motions for similar goal configurations. The use of PCA does not appear to improve the performance significantly, but it still helps to reduce the size of the data. In this case, for each path it reduces the number of variables from 30×14 ($D \times T$) to 50 (number of PCA components), more than 8 times reduction while maintaining the performance.

4.3 Planning from a random initial configuration to a random goal configuration

To proceed with a more complex case, the task here is to plan a path from a random initial configurations \mathbf{q}_{init} to a random target \mathbf{q}_{goal} . The task \mathbf{x} consists of the initial and goal configurations, $\mathbf{x} = (\mathbf{q}_{\text{init}}^\top, \mathbf{q}_{\text{goal}}^\top)^\top$. The database is constructed with $N_{\text{train}} = 500$ and evaluated with $N_{\text{test}} = 250$. The result is presented in Table 4.

k -NN performs poorly here, similar to STD, due to the dimension of the input space \mathbf{x} that is much larger as compared to Section 4.2. To achieve good performance, k -NN requires a much denser dataset. GPR outperforms BGMR by a wide margin.

Table 5: Planning from fixed \mathbf{q}_{init} to random Cartesian goal.

Method	Success (%)	Conv. time (s)	Cost (rad/s)
STD	65.6	0.98	1.849
k -NN	73.6	1.16	1.838
GPR	65.2	1.69	1.791
GPR _P CA	66.8	1.49	1.810
BGMR	72.4	1.30	1.755
BGMR _P CA	79.2	1.12	1.817
METRIC GPR _P CA	85.6	0.63	1.425

The last row of Table 4 shows the result of the ensemble method described in Section 3.2. Given an input \mathbf{x}^* , the method uses all function approximators to provide different warm-starts, each of which is used to initialize an instance of TrajOpt in parallel. Once a valid solution is obtained, the other instances of TrajOpt are terminated. This method results in a huge boost of the success rate, with comparable convergence time and cost to the other methods. If having a lower cost is more important than the computational time, the parallel execution can also be programmed such that we wait until all instances of TrajOpt finish the optimization, then the cost of the resulting paths are compared and the one with the lowest cost is chosen. This alternative implementation can reduce the cost, but at the expense of higher computational time.

4.4 Planning to Cartesian goals from a fixed initial configuration

In Section 4.2 and 4.3, we use TrajOpt to plan to goals in configuration space. In practical situations, however, the task is often to reach a certain Cartesian pose using the end-effector (e.g., to pick an object on the shelf), instead of planning to a specific joint configuration. One way to solve this problem is to first compute a configuration that achieves the Cartesian pose using inverse kinematics and plan to this configuration, but it does not make use of the flexibility inherent in the task. TrajOpt has an option to plan directly to a Cartesian goal, but it typically requires longer convergence time and lower success rate than planning to a joint configuration goal.

We present two approaches to use the memory of motion in this problem. In the first approach, we rely on the similar procedure as in previous cases: we formulate the task as $\mathbf{x} = (\mathbf{p}_{\text{left}}^\top, \mathbf{p}_{\text{right}}^\top)^\top$ where \mathbf{p}_{left} and $\mathbf{p}_{\text{right}}$ are the Cartesian positions of the right and left hand of PR2. The database is then constructed with $N_{\text{train}} = 1000$ and the function approximators are trained. In this approach, TrajOpt plans to a Cartesian goal directly. The second approach relies on the fact that a Cartesian goal corresponds to multiple goals in configuration space. In Section 4.2 we have already constructed a memory of motion that can predict an initial guess $\tilde{\mathbf{y}} = \tilde{\mathbf{q}}_{0:T}$, given a goal \mathbf{q}_{goal} in configuration space. The second approach uses this memory as a *metric* (Sect. 3.3) to choose between the different goals in configuration space. First, given a Cartesian goal \mathbf{x} , we run Inverse Kinematics (IK) solver to find $M = 5$ joint configurations that satisfy this pose. For each joint configuration, we use the memory of motion to predict the initial guess of the robot path to reach that configuration, and we compute the cost of each path. Finally, the goal configuration and the path with the lowest cost are chosen, and TrajOpt is run to reach this goal configuration with the given path as the warm-start. Note that in this approach, TrajOpt plans to a joint configuration instead of a Cartesian goal. For this approach we choose the method GPR_PCA from the Section 4.2, and use the term ‘METRIC GPR_PCA’ to differentiate from the first approach (denoted in standard notation).

We present the results in Table 5 with $N_{\text{test}} = 250$. Among the methods using the first approach, we note that BGMR yields better result than GPR because the mapping from the Cartesian goal \mathbf{x} to the robot path \mathbf{y} here is multimodal, as planning to a Cartesian pose has more redundancy as compared to planning to a joint configuration. This again demonstrates that BGMR handles multimodal output better than GPR. However, the second approach METRIC GPR_PCA outperforms even BGMR. The improvement over the first approach is very significant in all three criteria. This demonstrates that using the memory of motion as a metric to choose the optimal goal results in large improvements. We point out that the additional computational time required to find M IK solutions and the corresponding

warm-starts is only around 0.1 s, which is negligible compared to the convergence time.

5 Conclusion

We have presented an approach to build a memory of motion to warm-start trajectory optimization solver, and demonstrate through experiments with PR2 robot that the warm-start can improve the convergence time and the success rate while improving the quality of the solution. Function approximators and dimensionality reduction are used to learn the mapping between the task descriptor and the corresponding robot path. Three function approximators are considered: k -NN as baseline, GPR and BGMR. k -NN usually requires a dense dataset to obtain good performance, so it performs well when the size of the input space \mathbf{x} is small (as shown in Section 4.1, 4.2, and 4.4), while for larger input space (Section 4.4) it does not yield good results. GPR performs the best when the output is unimodal, while for multimodal output BGMR has a better performance than GPR. The use of PCA also improves the solution, although not too significantly, while reducing the memory storage. Using the ensemble method to warm-start the trajectory optimization boosts the success rate significantly. Finally, we have shown that the memory of motion can be used as a metric to choose optimally between several alternative goals, and this results in a significantly improved performance for the case of Cartesian goal planning.

Acknowledgment

This work was supported by the European Union under the EU H2020 collaborative project MEMMO (Memory of Motion, <http://www.memmo-project.eu/>), Grant Agreement No. 780684.

References

- [1] J. Schulman, J. Ho, A. X. Lee, I. Awwal, H. Bradlow, and P. Abbeel, “Finding locally optimal, collision-free trajectories with sequential convex optimization,” in *Proc. Robotics: Science and Systems (RSS)*, vol. 9, no. 1, 2013, pp. 1–10.
- [2] M. Toussaint, “Robot trajectory optimization using approximate inference,” in *Proc. Intl Conf. on Machine Learning (ICML)*. ACM, 2009, pp. 1049–1056.
- [3] M. Zucker, N. Ratliff, A. D. Dragan, M. Pivtoraiko, M. Klingensmith, C. M. Dellin, J. A. Bagnell, and S. S. Srinivasa, “Chomp: Covariant hamiltonian optimization for motion planning,” *Intl Journal of Robotics Research*, vol. 32, no. 9-10, pp. 1164–1193, 2013.
- [4] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal, “Stomp: Stochastic trajectory optimization for motion planning,” in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2011, pp. 4569–4574.
- [5] M. Stolle and C. G. Atkeson, “Policies based on trajectory libraries,” in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2006, pp. 3344–3349.
- [6] C. Liu and C. G. Atkeson, “Standing balance control using a trajectory library,” in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2009, pp. 3031–3036.
- [7] D. H. Jacobson and D. Q. Mayne, “Differential dynamic programming,” 1970.
- [8] N. Mansard, A. Del Prete, M. Geisert, S. Tonneau, and O. Stasse, “Using a memory of motion to efficiently warm-start a nonlinear predictive controller,” in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2018, pp. 2986–2993.
- [9] D. Forte, A. Gams, J. Morimoto, and A. Ude, “On-line motion synthesis and adaptation using a trajectory database,” *Robotics and Autonomous Systems*, vol. 60, no. 10, pp. 1327–1339, 2012.

- [10] A. J. Ijspeert, J. Nakanishi, and S. Schaal, “Learning rhythmic movements by demonstration using nonlinear oscillators,” in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2002, pp. 958–963.
- [11] L. E. Kavraki, P. Svestka, J. . Latombe, and M. H. Overmars, “Probabilistic roadmaps for robot path planning,” *IEEE Transactions on Robotics and Automation*, vol. 12, no. 4, pp. 566–580, 1996.
- [12] S. M. LaValle, “Rapidly-exploring random trees: A new tool for path planning,” Iowa State University, Computer Science Department, Tech. Rep. TR 98-11, 1998.
- [13] S. R. Martin, S. E. Wright, and J. W. Sheppard, “Offline and online evolutionary bi-directional rrt algorithms for efficient re-planning in dynamic environments,” in *Proc. IEEE Intl Conf. on Automation Science and Engineering (CASE)*, 2007, pp. 1131–1136.
- [14] D. Berenson, P. Abbeel, and K. Goldberg, “A robot path planning framework that learns from experience,” in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2012, pp. 3671–3678.
- [15] N. Jetchev and M. Toussaint, “Trajectory prediction: learning to map situations to robot trajectories,” in *Proc. Intl Conf. on Machine Learning (ICML)*, 2009, pp. 449–456.
- [16] W. Merkt, V. Ivan, and S. Vijayakumar, “Leveraging precomputation with problem encoding for warm-starting trajectory optimization in complex environments,” in *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, 2018, pp. 5877–5884.
- [17] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.
- [18] J. Quiñonero-Candela and C. E. Rasmussen, “A unifying view of sparse approximate gaussian process regression,” *Journal of Machine Learning Research*, vol. 6, no. Dec, pp. 1939–1959, 2005.
- [19] J. Hensman, N. Fusi, and N. D. Lawrence, “Gaussian processes for big data,” in *Conference on Uncertainty in Artificial Intelligence*, 2013, pp. 282–290.
- [20] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [21] S. Calinon, “A tutorial on task-parameterized movement learning and retrieval,” *Intelligent Service Robotics*, vol. 9, no. 1, pp. 1–29, 2016.
- [22] E. Pignat and S. Calinon, “Bayesian gaussian mixture model for robotic policy imitation,” *arXiv preprint arXiv:1904.10716*, 2019.
- [23] C. Doersch, “Tutorial on variational autoencoders,” *arXiv preprint arXiv:1606.05908*, 2016.
- [24] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.
- [25] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [26] F. Suárez-Ruiz, T. S. Lembono, and Q.-C. Pham, “Robotsp—a fast solution to the robotic task sequencing problem,” in *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2018, pp. 1611–1616.

A Supplementary Materials

A.1 Bayesian Gaussian Mixture Regression (BGMR)

As discussed in the main text, given the training dataset $\{\mathbf{X}, \mathbf{Y}\}$ GMR constructs the joint probability of (\mathbf{x}, \mathbf{y}) as a mixture of Gaussians,

$$p(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^K \pi_k \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (6)$$

where π_k , $\boldsymbol{\mu}_k$, and $\boldsymbol{\Sigma}_k$ are the k -th component's mixing coefficient, mean, and covariance, respectively. Let $\boldsymbol{\theta} = \{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}_{k=1}^K$, denoting the GMR parameters to be determined from the data.

We can decompose $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ according to \mathbf{x} and \mathbf{y} as

$$\boldsymbol{\mu}_k = \begin{pmatrix} \boldsymbol{\mu}_{k,x} \\ \boldsymbol{\mu}_{k,y} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}_k = \begin{pmatrix} \boldsymbol{\Sigma}_{k,xx} & \boldsymbol{\Sigma}_{k,xy} \\ \boldsymbol{\Sigma}_{k,yx} & \boldsymbol{\Sigma}_{k,yy} \end{pmatrix}. \quad (7)$$

Given a query \mathbf{x}^* , the predictive distribution of \mathbf{y} can then be computed by conditioning on \mathbf{x}^* ,

$$p(\mathbf{y} | \mathbf{x}^*, \boldsymbol{\theta}) = \sum_{k=1}^K p(k | \mathbf{x}^*, \boldsymbol{\theta}) p(\mathbf{y} | k, \mathbf{x}^*, \boldsymbol{\theta}), \quad (8)$$

where $p(k | \mathbf{x}^*, \boldsymbol{\theta})$ is the probability of \mathbf{x}^* belonging to the k -th component,

$$p(k | \mathbf{x}^*, \boldsymbol{\theta}) = \frac{\pi_k \mathcal{N}(\mathbf{x}^* | \boldsymbol{\mu}_{k,x}, \boldsymbol{\Sigma}_{k,xx})}{\sum_{i=1}^K \pi_i \mathcal{N}(\mathbf{x}^* | \boldsymbol{\mu}_{i,x}, \boldsymbol{\Sigma}_{i,xx})}, \quad (9)$$

and $p(\mathbf{y} | k, \mathbf{x}^*, \boldsymbol{\theta})$ is the predictive distribution of \mathbf{y} according to the k -th component,

$$p(\mathbf{y} | k, \mathbf{x}^*, \boldsymbol{\theta}) = \mathcal{N}\left(\boldsymbol{\mu}_{k,y} + \boldsymbol{\Sigma}_{k,yx} \boldsymbol{\Sigma}_{k,xx}^{-1} (\mathbf{x}^* - \boldsymbol{\mu}_{k,x}), \boldsymbol{\Sigma}_{k,yy} - \boldsymbol{\Sigma}_{k,yx} \boldsymbol{\Sigma}_{k,xx}^{-1} \boldsymbol{\Sigma}_{k,xy}\right), \quad (10)$$

which is a Gaussian distribution with the mean being linear in \mathbf{x}^* . The resulting predictive distribution (8) is then a mixture of Gaussians.

In GMR, $\boldsymbol{\theta}$ is usually estimated from the dataset $\{\mathbf{X}, \mathbf{Y}\}$ by using Expectation-Minimization method that outputs a single value for $\boldsymbol{\theta}$, and the prediction is computed using this value. Instead, Bayesian GMR (BGMR) puts priors to the parameter $\boldsymbol{\theta}$, and given the dataset $\{\mathbf{X}, \mathbf{Y}\}$, it computes the posterior distribution $p(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})$. Hence, when we do prediction we have to integrate over the posterior distribution of $\boldsymbol{\theta}$,

$$p(\mathbf{y} | \mathbf{x}^*, \mathbf{X}, \mathbf{Y}) = \int_{\boldsymbol{\theta}} p(\mathbf{y} | \mathbf{x}^*, \boldsymbol{\theta}) p(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y}) d\boldsymbol{\theta}. \quad (11)$$

The resulting predictive distribution is then a mixture of multivariate t-distributions,

$$p(\mathbf{y} | \mathbf{x}^*, \mathbf{X}, \mathbf{Y}) = \sum_{k=1}^K p(k | \mathbf{x}^*, \mathbf{X}, \mathbf{Y}) p(\mathbf{y} | k, \mathbf{x}^*, \mathbf{X}, \mathbf{Y}), \quad (12)$$

where $p(k | \mathbf{x}^*, \mathbf{X}, \mathbf{Y})$ is the probability of \mathbf{x}^* belonging to the k -th component and $p(\mathbf{y} | k, \mathbf{x}^*, \mathbf{X}, \mathbf{Y})$ is the predictive distribution of \mathbf{y} according to the k -th component. These are calculated using the formula similar to (9) and (10), but the Gaussian distributions are replaced by t-distributions. $p(\mathbf{y} | k, \mathbf{x}^*, \mathbf{X}, \mathbf{Y})$ is a t-distribution (with the mean being linear w.r.t. \mathbf{x}^*) instead of a Gaussian distribution as in (10), due to the integration over the posterior distribution in (11).

Note that the predictive distribution of GMR (8) and BGMR (12) are very similar, but (8) is a mixture of Gaussians whereas (12) is a mixture of t-distributions. Both can be seen as a mixture of linear regressions, since the mean prediction according to each k -th component (10) is linear in \mathbf{x}^* .

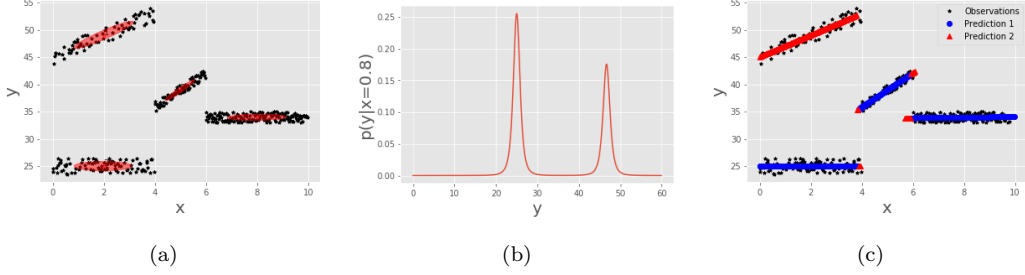


Figure 3: Toy example of multimodal and discontinuous mapping $f : x \rightarrow y$. BGMR fits a mixture of Gaussians to the joint data of \mathbf{x} and \mathbf{y} (a), resulting in a mixture of t-distributions as the predictive distribution of y given $x^* = 0.8$ (b). By taking the means of the highest and second highest component of the mixture, BGMR approximates the mapping very well (c).

The advantage of having a Bayesian version is GMR is the following:

- When the dimension of \mathbf{x} and \mathbf{y} is high, the number of parameters to be estimated in $\boldsymbol{\theta}$ is large and the estimation of $\boldsymbol{\theta}$ in GMR is prone to overfitting, unless the number of training data is very large. Putting the priors to $\boldsymbol{\theta}$ in BGMR helps mitigation of overfitting and makes it possible to learn with fewer number of datapoints.
- By putting Dirichlet Process prior to the parameters π_k , the number of components K can be determined automatically.

More details about the priors and the derivations of BGMR can be found in [22].

To illustrate how BGMR can handle multimodal output, we show a toy example in Figure 3, which depicts a mapping $y = f(x)$ that is discontinuous at $x = 4$ and $x = 6$, and has two modes at $0 < x < 4$. The data points corresponding to (x, y) are shown in black. First, BGMR fits a mixture of Gaussians to $\{x, y\}$, shown as red ellipses in Figure 3a. We can see that the Gaussians fit well each mode and each side of discontinuity.

Given a query point x^* , the predictive distribution of y^* is a mixture of t-distributions given by (12). Figure 3b shows an example of $p(y|x)$ for $x = 0.8$. We can see that there are two peaks corresponding to the two modes of the mapping f (i.e., $y = 25.2$ and $y = 46.7$), having roughly the same probability (0.5).

To have a single-point prediction x^* from the predictive distribution, we can take the mean of the component that has the highest probability. Figure 3c shows the BGMR predictions over the range of x , by taking the mean of the components with the highest and the second-highest probability in (12). We see that it predicts both modes very well, and it still performs well at discontinuous region (i.e., it does not average the both sides of discontinuity).

A.2 Algorithms

We present here the pseudocode for the methods discussed in Section 3.

Algorithm 1 Building Memory of Motion

INPUT: number of samples N
OUTPUT: the database $\{\mathbf{X}, \mathbf{Y}\}$ and the function approximator \mathbf{f}

- 1: $\mathbf{X} = [], \mathbf{Y} = []$
- 2: **for** $i = 1, 2, \dots, N$ **do**
- 3: sample a random task \mathbf{x}_i
- 4: compute the initial guess $\tilde{\mathbf{y}}_i$ to achieve \mathbf{x}_i by straight-line motion
- 5: solve \mathbf{x}_i using TrajOpt warm-started by $\tilde{\mathbf{y}}_i$, to obtain the path \mathbf{y}_i
- 6: **if** \mathbf{y}_i is valid **then**
- 7: add $(\mathbf{x}_i, \mathbf{y}_i)$ to $\{\mathbf{X}, \mathbf{Y}\}$
- 8: **end if**
- 9: **end for**
- 10: apply PCA to \mathbf{Y} to obtain $\hat{\mathbf{Y}}$ (*Optional*)
- 11: train the function approximator \mathbf{f} on $\{\mathbf{X}, \mathbf{Y}\}$ (*or on $\{\mathbf{X}, \hat{\mathbf{Y}}\}$ if PCA is used*)

Algorithm 2 Ensemble Method

INPUT: Task \mathbf{x}^* , a list of function approximators $\{\mathbf{f}_j\}_{j=1}^M$
OUTPUT: The path \mathbf{y}^* that accomplishes the task \mathbf{x}^*

- 1: **for all** $j = 1, 2, \dots, M$ **do in parallel**
- 2: compute the initial guess $\tilde{\mathbf{y}}_j = \mathbf{f}_j(\mathbf{x}^*)$
- 3: solve \mathbf{x}^* using TrajOpt warm-started by $\tilde{\mathbf{y}}_j$, to obtain the path \mathbf{y}_j
- 4: **if** \mathbf{y}_j is valid **then**
- 5: $\mathbf{y}^* = \mathbf{y}_j$
- 6: Terminate the parallel execution
- 7: **end if**
- 8: **end for**

Algorithm 3 Using Memory of Motion as Metric

INPUT: A list of goals $\{\mathbf{x}_j\}_{j=1}^M$, a function approximator \mathbf{f}
OUTPUT: The optimal goal \mathbf{x}^* and the corresponding path \mathbf{y}^*

- 1: **for** $j = 1, 2, \dots, M$ **do**
- 2: compute the initial guess $\tilde{\mathbf{y}}_j = \mathbf{f}(\mathbf{x}_j)$
- 3: compute the cost $\ell(\tilde{\mathbf{y}}_j)$
- 4: **end for**
- 5: $j^* \leftarrow \arg \min_j \ell(\tilde{\mathbf{y}}_j)$
- 6: $\mathbf{x}^* \leftarrow \mathbf{x}_{j^*}$
- 7: $\tilde{\mathbf{y}}^* \leftarrow \tilde{\mathbf{y}}_{j^*}$
- 8: solve \mathbf{x}^* using TrajOpt warm-started by $\tilde{\mathbf{y}}^*$, to obtain the path \mathbf{y}^*
