

# Learning Manifolds for Sequential Motion Planning

Isabel M. Rayas Fernández\*, Giovanni Sutanto\*, Peter Englert, Ragesh K. Ramachandran, Gaurav S. Sukhatme

**Abstract**—Motion planning with constraints is an important part of many real-world robotic systems. In this work, we study manifold learning methods to learn such constraints from data. We explore two methods for learning implicit constraint manifolds from data: Variational Autoencoders (VAE), and a new method, Equality Constraint Manifold Neural Network (ECoMaNN). With the aim of incorporating learned constraints into a sampling-based motion planning framework, we evaluate the approaches on their ability to learn representations of constraints from various datasets and on the quality of paths produced during planning.

## I. INTRODUCTION

Many practical robotic applications require planning robot motions with constraints, such as maintaining an orientation or reaching a particular location. Planning becomes more complicated when the task consists of many subtasks that must be completed in sequence. In this case, task and motion planning frameworks [11, 6, 16, 23, 7, 1, 3] can be used to handle long planning horizons and a wide range of tasks. However, some constraints may be difficult to describe analytically, or it may be difficult to sample constraints that adhere to them. For example, if the task is to pour a liquid from a bottle into a cup, it is not immediately clear how to encode the motion constraints for a planning algorithm.

In this work, we focus on learning constraint manifolds for use in constrained motion planning algorithms [9, 19, 2, 10, 13, 15, 21, 5]. To this end, we investigate two different approaches: Variational autoencoders (VAE) following Kingma and Welling [14], and Equality Constraint Manifold Neural Network (ECoMaNN), a method we propose which learns the implicit function value of equality constraints. We evaluate these techniques on six datasets of varying size and complexity, and we present preliminary resulting motion plans.

## II. BACKGROUND ON SEQUENTIAL MANIFOLD PLANNING

We focus on learning manifolds that describe kinematic robot tasks. We aim to integrate these learned manifolds into the sequential manifold planning (SMP) framework proposed in Englert et al. [9]. SMP considers kinematic motion planning problems in a configuration space  $\mathcal{C} \subseteq \mathbb{R}^d$ . A robot configuration  $\mathbf{q} \in \mathcal{C}$  describes the state of one or more robots with  $d$  degrees of freedom in total. A manifold  $M$  is represented as an equality constraint  $h_M(\mathbf{q}) = \mathbf{0}$  where

$h_M : \mathbb{R}^d \rightarrow \mathbb{R}^l$  and  $l$  is the dimensionality of the implicit manifold. The set of robot configurations that are on a manifold  $M$  is given by  $\mathcal{C}_M = \{\mathbf{q} \in \mathcal{C} \mid h_M(\mathbf{q}) = \mathbf{0}\}$ . SMP defines the planning problem as a sequence of  $(n + 1)$  such manifolds  $\mathcal{M} = \{M_1, M_2, \dots, M_{n+1}\}$  and an initial configuration  $\mathbf{q}_{\text{start}} \in \mathcal{C}_{M_1}$  that is on the first manifold. The goal is to find a path from  $\mathbf{q}_{\text{start}}$  that traverses the manifold sequence  $\mathcal{M}$  and reaches a configuration on the goal manifold  $M_{n+1}$ . A path on the  $i$ -th manifold is defined as  $\tau_i : [0, 1] \rightarrow \mathcal{C}_{M_i}$  and  $J(\tau_i)$  is the cost function of a path  $\mathcal{J} : \mathcal{T} \rightarrow \mathbb{R}_{\geq 0}$  where  $\mathcal{T}$  is the set of all non-trivial paths. The problem is formulated as an optimization over a set of paths  $\tau = (\tau_1, \dots, \tau_n)$  that minimizes the sum of path costs under the constraints of traversing  $\mathcal{M}$  and of being collision-free:

$$\begin{aligned} \tau^* &= \arg \min_{\tau} \sum_{i=1}^n J(\tau_i) \\ \text{s.t. } \tau_1(0) &= \mathbf{q}_{\text{start}} \\ \tau_i(1) &= \tau_{i+1}(0) && \forall_{i=1, \dots, n-1} \\ \mathcal{C}_{\text{free}, i+1} &= \Upsilon(\mathcal{C}_{\text{free}, i}, \tau_i) && \forall_{i=1, \dots, n} \\ \tau_i(s) &\in \mathcal{C}_{M_i} \cap \mathcal{C}_{\text{free}, i} && \forall_{i=1, \dots, n} \forall_{s \in [0, 1]} \\ \tau_n(1) &\in \mathcal{C}_{M_{n+1}} \cap \mathcal{C}_{\text{free}, n+1} \end{aligned} \quad (1)$$

$\Upsilon$  is an operator that describes the change in the free configuration space (the space of all configurations that are not in collision with the environment)  $\mathcal{C}_{\text{free}}$  when transitioning to the next manifold. The SMP algorithm is able to solve this problem for a certain class of problems. It iteratively applies RRT\* to find a path that reaches the next manifold while staying on the current manifold. For further details of the SMP algorithm, we refer the reader to Englert et al. [9].

In this paper, we employ data-driven algorithms to learn manifolds  $M$  from data with the goal to integrate them into the SMP framework.

## III. MANIFOLD LEARNING

Learning constraint manifolds from data is attractive for multiple reasons. For example, it may be easier for a human to demonstrate a task rather than specifying constraints analytically, or we may want to reduce the amount of expert information needed.

We propose a novel neural network structure – called *Equality Constraint Manifold Neural Network* (ECoMaNN) – to become a learning representation that takes  $\mathbf{q}$  as input and outputs the prediction of the implicit function  $h_M(\mathbf{q})$ . Moreover, we would like to train ECoMaNN in a supervised manner, from demonstrations. One of the challenges is that the supervised training dataset is collected *only* from demonstrations of data points which are on the equality constraint manifold  $\mathcal{C}_M$ , called the *on-manifold* dataset. This is

\* equal contribution

All authors are with the Robotic Embedded Systems Laboratory, University of Southern California, Los Angeles, CA, USA.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1842487. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work was supported in part by the Office of Naval Research (ONR) under grant N000141512550.

a reasonable assumption, since collecting both the on-manifold  $\mathcal{C}_M$  and off-manifold  $\mathcal{C}_{\bar{M}} = \{\mathbf{q} \in \mathcal{C} \mid h_M(\mathbf{q}) \neq \mathbf{0}\}$  datasets for supervised training will be tedious because the implicit function  $h_M$  values of points in  $\mathcal{C}_{\bar{M}}$  are typically unknown and hard to label. We will show that even though our approach is only provided with data on  $\mathcal{C}_M$ , it can still learn a useful representation of the manifold, sufficient for use in the SMP framework.

Our goal is to learn a single global representation of the constraint manifold in form of a neural network. A manifold can be defined as a collection of local neighborhoods which resemble Euclidean spaces (Lee [17]). Therefore, a global representation of the manifold can be developed by constructing characterizations for its Euclidean-like local neighborhoods.

Our approach leverages local information on the manifold in the form of the tangent and normal spaces (Deutsch and Medioni [8], Strang [20]). With regard to  $h_M$ , the tangent and normal spaces are equivalent to the null and row space, respectively, of the matrix  $\mathbf{J}_M = \left. \frac{\partial h_M(\mathbf{q})}{\partial \mathbf{q}} \right|_{\mathbf{q}=\bar{\mathbf{q}}}$ , and valid in a small neighborhood around the point  $\bar{\mathbf{q}}$ .

Using on-manifold data, the local information of the manifold can be analyzed using Local Principal Component Analysis (Local PCA) (Kambhatla and Leen [12]). Essentially, for each data point  $\mathbf{q}$  in the on-manifold dataset, we establish a local neighborhood using  $K$ -nearest neighbors ( $KNN$ )  $\hat{\mathcal{K}} = \{\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2, \dots, \hat{\mathbf{q}}_K\}$ , with  $K \geq d$ . After a change of coordinates,  $\mathbf{q}$  becomes the origin of a new local coordinate frame  $\mathcal{F}$ , and the  $KNN$  becomes  $\tilde{\mathcal{K}} = \{\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2, \dots, \tilde{\mathbf{q}}_K\}$  with  $\tilde{\mathbf{q}}_k = \hat{\mathbf{q}}_k - \mathbf{q}$  for all values of  $k$ . Defining the matrix  $\mathbf{X} = [\tilde{\mathbf{q}}_1 \ \tilde{\mathbf{q}}_2 \ \dots \ \tilde{\mathbf{q}}_K]^T \in \mathbb{R}^{K \times d}$ , we can compute the covariance matrix  $\mathbf{S} = \frac{1}{K-1} \mathbf{X}^T \mathbf{X} \in \mathbb{R}^{d \times d}$ . The eigendecomposition of  $\mathbf{S} = \mathbf{V} \mathbf{\Sigma} \mathbf{V}^T$  gives us the Local PCA. The matrix  $\mathbf{V}$  contains the eigenvectors of  $\mathbf{S}$  as its columns in decreasing order w.r.t. the corresponding eigenvalues in the diagonal matrix  $\mathbf{\Sigma}$ . These eigenvectors form the basis of  $\mathcal{F}$ .

This local coordinate frame  $\mathcal{F}$  is tightly related to the tangent and normal spaces of the manifold at  $\mathbf{q}$ . That is, the  $(d-l)$  eigenvectors corresponding to the  $(d-l)$  biggest eigenvalues of  $\mathbf{\Sigma}$  form a basis of the tangent space, while the remaining  $l$  eigenvectors form the basis of the normal space. Furthermore, due to the characteristics of the manifold from which the dataset was collected, the  $l$  smallest eigenvalues of  $\mathbf{\Sigma}$  will be close to zero, resulting in the  $l$  eigenvectors associated with them forming the basis of the null space of  $\mathbf{S}$ . On the other hand, the remaining  $(d-l)$  eigenvectors form the basis of the row space of  $\mathbf{S}$ .

To this end, we present several methods to define and train ECoMaNN, as follows:

#### A. Local Tangent and Normal Spaces Alignment

ECoMaNN aims to align the following:

- the null space of  $\mathbf{J}_M$  and the row space of  $\mathbf{S}$ , which both must be equivalent to the tangent space, and
- the row space of  $\mathbf{J}_M$  and the null space of  $\mathbf{S}$ , which both must be equivalent to the normal space

for each local neighborhood of each point  $\mathbf{q}$  in the on-manifold dataset. Suppose the eigenvectors of  $\mathbf{S}$  are  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d\}$

and the singular vectors of  $\mathbf{J}_M$  are  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d\}$ , where the indices indicate the decreasing order w.r.t. the eigenvalue/singular value magnitude. The null spaces of  $\mathbf{S}$  and  $\mathbf{J}_M$  are spanned by  $\{\mathbf{v}_{d-l+1}, \dots, \mathbf{v}_d\}$  and  $\{\mathbf{e}_{l+1}, \dots, \mathbf{e}_d\}$ , respectively. The two conditions above imply that the projection of the null space eigenvectors of  $\mathbf{J}_M$  into the null space of  $\mathbf{S}$  should be  $\mathbf{0}$ , and similarly for the row spaces. Hence, we achieve this by training ECoMaNN to minimize projection errors  $\|\mathbf{V}_N \mathbf{V}_N^T \mathbf{E}_N\|_2^2$  and  $\|\mathbf{E}_N \mathbf{E}_N^T \mathbf{V}_N\|_2^2$  with  $\mathbf{V}_N = [\mathbf{v}_{d-l+1} \ \dots \ \mathbf{v}_d]$  and  $\mathbf{E}_N = [\mathbf{e}_{l+1} \ \dots \ \mathbf{e}_d]$ .

#### B. Data Augmentation with Off-Manifold Data

The training dataset is on-manifold, i.e., each point  $\mathbf{q}$  in the dataset satisfies  $h_M(\mathbf{q}) = \mathbf{0}$ . Through Local PCA on each of these points, we know the data-driven approximation of the normal space of the manifold at  $\mathbf{q}$ . Hence, we know the directions where the violation of the equality constraint increases, i.e., the same or opposite direction of any vector from the approximate normal space. Since our future use of the learned constraint manifold on motion planning does not require the acquisition of the near-ground-truth value of  $h_M(\mathbf{q}) \neq \mathbf{0}$ , we can set this off-manifold valuation of  $h_M$  arbitrarily, as long as it does not interfere with the utility for projecting an off-manifold point onto the manifold. Therefore, we can augment our dataset with additional off-manifold data to achieve a more robust learning of ECoMaNN. For each point  $\mathbf{q}$  in the on-manifold dataset, and for each eigenvector  $\mathbf{v}$  that forms the basis of the normal space at  $\mathbf{q}$ , we can add an off-manifold point  $\tilde{\mathbf{q}} = \mathbf{q} + i\epsilon\mathbf{v}$  with a non-zero signed integer  $i$  and a small positive scalar  $\epsilon$ . For such an augmented data point  $\tilde{\mathbf{q}}$ , we set the label satisfying  $\|h_M(\tilde{\mathbf{q}})\|_2 = |i|\epsilon$ . During training, we minimize the prediction error  $\|(\|h_M(\tilde{\mathbf{q}})\|_2 - |i|\epsilon)\|_2^2$  for each augmented point  $\tilde{\mathbf{q}}$ .

## IV. DATASETS

We use a robot simulator (Todorov et al. [22]) to generate various datasets. For each dataset, we define  $h_M(\mathbf{q})$  by hand and randomly sample points in the configuration space and project them onto the manifold. We use six datasets:

- **Nav:** 2D point that has to stay close to a reference point. Defined as an inequality constraint.  $N = 15000$ .
- **Sphere:** 3D point that has to stay on the surface of a sphere.  $N = 10000$ .
- **Plane:** Robot arm with 3 rotational DOFs where the end effector has to be on a plane.  $N = 999$ .
- **Orient:** Robot arm with 6 rotational DOFs that has to keep its orientation upright (e.g., transporting a cup).  $N = 21153$ .
- **Tilt:** Same as Orient, but here the orientation constraint is relaxed to an inequality constraint.  $N = 2000$ .
- **Handover:** Robot arm with 6 rotational DOFs and a mobile base with 2 translational DOFs. The manifold is defined as an equality constraint that describes the handover of an object between the two robots.  $N = 2002$ .

## V. EXPERIMENTS

We compare the proposed ECoMaNN method to a Variational Autoencoder (VAE), which is a popular method for learning a generative model of a set of data (Chen et al. [4], Kingma and Welling [14], Park et al. [18]). Importantly, because they embed data points as a distribution in the latent space, new latent vectors can be sampled and decoded into unseen examples which fit the distribution of the training data. VAEs make use of two neural networks in a neural autoencoder structure during training, and they only use the decoder during generation. The key idea that makes VAEs computationally tractable is that the distribution in the latent space is assumed to be Gaussian. The loss function is a combination of the reconstruction error of the input and the KL divergence of the latent space distribution, weighted by a parameter  $\beta$ .

We use the following network structures and parameters: For the Nav dataset, the VAE has two hidden layers with 6 and 4 units. The input size is 2 and the embedding size is 2. For the Plane dataset, the VAE has three hidden layers with 12, 9, and 6 units. The input size is 3 and the embedding size is 2. For the Sphere, Orient, Tilt, and Handover datasets, the VAEs have the same structure: Four hidden layers with 8, 6, 6, and 4 units. The input sizes to the networks are 3, 6, 6, and 8, and the embedding sizes are 2, 5, 3, and 7, respectively. All VAE models have  $\beta = 0.25$  and use batch normalization. We train for 500 epochs for Handover, and 200 otherwise.

### A. Evaluate Implicit Functions on Datasets

We compare the performance of the models using the implicit function value  $h_M$ . In the case of the VAE models, we take the reprojected data  $\hat{X}$  and evaluate each configuration with  $h_M$ . In the case of the ECoMaNN, the output of the network is the estimated implicit function value of the input, so we can directly use it. We report the mean and standard deviation of  $h_M$  for each dataset in Table I. Note that for Nav and Tilt datasets,  $h_M$  does not need to be 0 for a configuration to be valid, since these are inequality constraints. Values less than 1 for Nav and less than 0.1 for Tilt adhere to the constraints.

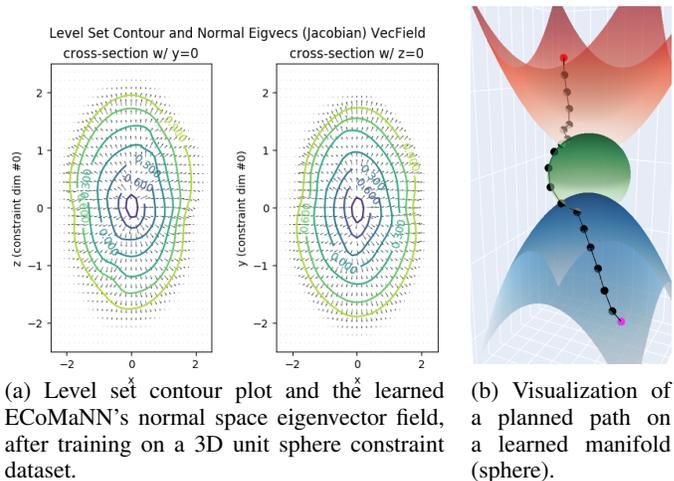
In Fig. 1a, we plot the level set contour as well as the normal space eigenvector field of an ECoMaNN after training on a 3D unit sphere constraint dataset. We see that at both cross-sections  $y = 0$  (left) and  $z = 0$  (right), the contours are close to circular, which is expected for a unit sphere constraint manifold.

### B. Evaluate Implicit Functions on Generated Configurations

We use the trained VAE models from Sec. V-A to generate 100,000 new, on-manifold configurations for each constraint. We then evaluate these configurations with the implicit function  $h_M$  and report the mean and standard deviation in Table II.

### C. Sequential Motion Planning on Learned Manifolds

In this experiment, we incorporate a learned manifold into the planning framework developed and introduced in Englert



(a) Level set contour plot and the learned ECoMaNN's normal space eigenvector field, after training on a 3D unit sphere constraint dataset. (b) Visualization of a planned path on a learned manifold (sphere).

TABLE I: Experiment A –  $h_M$  on Datasets.

Method	Dataset											
	Nav		Sphere		Plane		Orient		Tilt		Handover	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
VAE	0.42	0.18	0.08	0.03	0.16	0.12	0.01	0.03	0.09	0.07	0.56	0.47
ECoMaNN	N/A	N/A	0.04	0.06	0.46	1.39	-0.05	0.93	N/A	N/A	0.24	0.76

et al. [9]. The Sphere dataset (see Section IV) is used to learn a manifold representation with ECoMaNN. This learned manifold is combined with two analytical manifolds representing paraboloids. A motion planning problem is defined on these three manifolds where a 3D point has to find a path from a start configuration on one of the paraboloids to a goal configuration on the other. See Figure 1b for a visualization of the manifolds and a found path with SMP.

## VI. DISCUSSION AND FUTURE WORK

In this paper, we presented ways of learning constraint manifolds for sequential manifold planning. One of them is the novel Equality Constraint Manifold Neural Network (ECoMaNN). ECoMaNN is a method for learning representation for implicit functions, with an emphasis on representing equality constraints, while VAEs can also learn inequality constraints. We showed that ECoMaNN has successfully learned equality constraint manifolds and that these manifolds can be used in a sequential motion planning method.

There are several interesting improvements and future directions to pursue. First, there are still limitations with the current approach; in particular, our approach does not address the sign/polarity assignments of the implicit function value output, which we plan to address. Second, we plan to do more comprehensive testing on higher-dimensional manifolds, and incorporate multiple learned constraints into a single motion plan with more subtasks. Further, we also plan to integrate manifolds learned with VAE into motion planning algorithms.

TABLE II: Experiment B –  $h_M$  on Generated Configurations.

Method	Dataset											
	Nav		Sphere		Plane		Orient		Tilt		Handover	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
VAE	0.58	0.18	0.10	0.03	0.14	0.08	0.01	0.01	0.96	0.30	0.10	0.02

## REFERENCES

- [1] Jennifer Barry, Leslie Pack Kaelbling, and Tomás Lozano-Pérez. A hierarchical approach to manipulation with diverse actions. In *2013 IEEE International Conference on Robotics and Automation*, pages 1799–1806. IEEE, 2013.
- [2] Dmitry Berenson, Siddhartha Srinivasa, and James Kuffner. Task space regions: A framework for pose-constrained manipulation planning. *The International Journal of Robotics Research*, 30(12):1435–1460, 2011.
- [3] Stéphane Cambon, Rachid Alami, and Fabien Gravot. A hybrid approach to intricate motion, manipulation and task planning. *The International Journal of Robotics Research*, 28(1):104–126, 2009.
- [4] N. Chen, M. Karl, and P. van der Smagt. Dynamic movement primitives in latent space of time-dependent variational autoencoders. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 629–636, 2016.
- [5] Juan Cortés and Thierry Siméon. Sampling-based motion planning under kinematic loop-closure constraints. In *Algorithmic Foundations of Robotics VI*, pages 75–90. Springer, 2004.
- [6] Neil T. Dantam, Zachary K. Kingston, Swarat Chaudhuri, and Lydia E. Kavraki. Incremental Task and Motion Planning: A Constraint-Based Approach. In *Proceedings of Robotics: Science and Systems*, 2016.
- [7] Neil T Dantam, Swarat Chaudhuri, and Lydia E Kavraki. The Task-Motion Kit: An open source, general-purpose task and motion-planning framework. *IEEE Robotics & Automation Magazine*, 25(3):61–70, 2018.
- [8] Shay Deutsch and Gérard Medioni. Unsupervised learning using the tensor voting graph. In Jean-François Aujol, Mila Nikolova, and Nicolas Papadakis, editors, *Scale Space and Variational Methods in Computer Vision*, pages 282–293, Cham, 2015. Springer International Publishing. ISBN 978-3-319-18461-6.
- [9] Peter Englert, Isabel M. Rayas Fernández, Ragesh K. Ramachandran, and Gaurav S. Sukhatme. Sampling-Based Motion Planning on Manifold Sequences. arXiv:2006.02027, 2020.
- [10] Léonard Jaillet and Josep M Porta. Asymptotically-optimal path planning on manifolds. In *Proceedings of Robotics: Science and Systems*, 2013.
- [11] Leslie Pack Kaelbling and Tomás Lozano-Pérez. Integrated task and motion planning in belief space. *The International Journal of Robotics Research*, 32(9-10):1194–1227, 2013.
- [12] Nandakishore Kambhatla and Todd K. Leen. Dimension Reduction by Local Principal Component Analysis. *Neural Comput.*, 9(7):1493–1516, October 1997. ISSN 0899-7667. doi: 10.1162/neco.1997.9.7.1493. URL <https://doi.org/10.1162/neco.1997.9.7.1493>.
- [13] Beobkyoon Kim, Terry Taewoong Um, Chansu Suh, and Frank C Park. Tangent bundle RRT: A randomized algorithm for constrained motion planning. *Robotica*, 34(1):202–225, 2016.
- [14] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. *CoRR*, abs/1312.6114, 2014.
- [15] Zachary Kingston, Mark Moll, and Lydia E Kavraki. Exploring implicit spaces for constrained sampling-based planning. *The International Journal of Robotics Research*, 38(10-11):1151–1178, 2019.
- [16] George Konidaris, Leslie Pack Kaelbling, and Tomas Lozano-Perez. From skills to symbols: Learning symbolic representations for abstract high-level planning. *Journal of Artificial Intelligence Research*, 61:215–289, 2018.
- [17] John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2006.
- [18] D. Park, Y. Hoshi, and C. C. Kemp. A Multimodal Anomaly Detector for Robot-Assisted Feeding Using an LSTM-Based Variational Autoencoder. *IEEE Robotics and Automation Letters*, 3(3):1544–1551, 2018.
- [19] Mike Stilman. Global manipulation planning in robot joint space with task constraints. *IEEE Transactions on Robotics*, 26(3):576–584, 2010.
- [20] Gilbert Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, Wellesley, MA, fourth edition, 2009. ISBN 9780980232714 0980232716 9780980232721 0980232724 9788175968110 8175968117.
- [21] Ioan A Şucan and Sachin Chitta. Motion planning with constraints using configuration space approximations. In *Proceedings of the International Conference on Intelligent Robots and Systems*, 2012.
- [22] E. Todorov, T. Erez, and Y. Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033, 2012.
- [23] Marc Toussaint, Kelsey Allen, Kevin A Smith, and Joshua B Tenenbaum. Differentiable Physics and Stable Modes for Tool-Use and Manipulation Planning. In *Proceedings of Robotics: Science and Systems*, 2018.