

TamedPUMA: safe and stable imitation learning with geometric fabrics

Saray Bakker

Department of Mechanical Engineering, TU Delft, The Netherlands

S.BAKKER-7@TUDELFT.NL

Rodrigo Pérez-Dattari

Department of Mechanical Engineering, TU Delft, The Netherlands

R.J.PEREZDATTARI@TUDELFT.NL

Cosimo Della Santina

Department of Mechanical Engineering, TU Delft, The Netherlands

C.DELLASANTINA@TUDELFT.NL

Wendelin Böhmer

Department of Electrical Engineering, Mathematics & Computer Science, TU Delft, The Netherlands

J.W.BOHMER@TUDELFT.NL

Javier Alonso-Mora

Department of Mechanical Engineering, TU Delft, The Netherlands

J.ALONSOMORA@TUDELFT.NL

Editors: N. Ozay, L. Balzano, D. Panagou, A. Abate

Abstract

Using the language of dynamical systems, Imitation learning (IL) provides an intuitive and effective way of teaching stable task-space motions to robots with goal convergence. Yet, IL techniques are affected by serious limitations when it comes to ensuring safety and fulfillment of physical constraints. With this work, we solve this challenge via TamedPUMA, an IL algorithm augmented with a recent development in motion generation called geometric fabrics. As both the IL policy and geometric fabrics describe motions as artificial second-order dynamical systems, we propose two variations where IL provides a navigation policy for geometric fabrics. The result is a stable imitation learning strategy within which we can seamlessly blend geometrical constraints like collision avoidance and joint limits. Beyond providing a theoretical analysis, we demonstrate TamedPUMA with simulated and real-world tasks, including a 7-DoF manipulator. [Link to website](#)

Keywords: Imitation Learning, Dynamical Systems, Geometric Motion Generation, Fabrics, Movement Primitives

1. Introduction

As robotic solutions rapidly enter unstructured environments such as the agriculture sector and homes, there is a critical need for methods that allow non-experts to easily adapt robots for new tasks. Currently, experts manually program these tasks, a method that is costly and not scalable for widespread use. Moreover, these sectors demand that robots safely interact with dynamic environments where humans are present.

A possible solution to this societal challenge comes from Imitation Learning (IL). Using this technique, robots can learn motion profiles from demonstrations provided by non-expert users. Furthermore, by encoding the learned trajectories as solutions of a dynamical system, established mathematical tools from dynamical system theory can be used to guarantee convergence to the task's goal state, as we discuss more in detail in Sec. 1.1. In a robotics context, these learned dynamical systems commonly encode the navigation policy towards a goal within a task space - such as the evolution of the end-effector pose of a manipulator while pouring water in a glass from all possible

initial locations. However, this focus on task space motions renders IL fundamentally limited when considering physical constraints involving the robot’s body impacting itself or interacting with the external environment. Substantial recent research has looked into the problem but, as we discuss in Sec. 1.1, to the best of our knowledge, no IL method can simultaneously ensure stability and real-time fulfillment of physical constraints for systems with many degrees of freedom.

This work contributes to the state of the art by introducing TamedPUMA, a learning framework that builds on IL and geometric fabrics (Ratliff et al., 2020) to obtain inherently stable and safe motion primitives from demonstrations. By leveraging the recently introduced geometric framework for motion generation called geometric fabrics, our approach learns stable motion profiles while considering online whole-body collision-avoidance and joint limits. To make this possible, the learned task-space policy has to be formulated as a 2nd-order (neural) dynamical system as fabrics operate within the Finsler Geometry framework where vector fields must be defined at the acceleration level (Bao et al., 2012). Also, the learned policy must admit an *aligned* potential, which roughly means a scalar function whose gradient is aligned with the acceleration field when the velocity is null. In this paper, we formally assess how we can ensure that both conditions are met. The performance of our approach is evaluated with a simulated and real-world 7-DoF manipulator, where we also benchmark it against vanilla geometric fabrics, vanilla learned stable motion primitives and a modulation-based IL approach leveraging collision-aware IK.

1.1. Related work

Learning stable dynamical systems from demonstrations Multiple methodologies have been introduced for learning stable dynamical systems (Billard et al., 2022; Hu et al., 2024). An influential methodology in learning stable dynamical systems is Dynamical Movement Primitives (DMP) by Ijspeert et al. (2013), ensuring convergence towards a simple manually-designed dynamical system. This approach is extended to non-Euclidean state spaces (Ude et al., 2014), probabilistic environments (Li et al., 2023; Przystupa et al., 2023), and in the context of deep neural network (DNN) (Pervez et al., 2017; Ridge et al., 2020). To ensure stability, several approaches enforce a specific structure on the function approximators, such as enforcing positive or negative definiteness (Khansari-Zadeh and Billard, 2011; Lemme et al., 2014; Fernandez and Billard, 2018), or invertibility (Perrin and Schlehuber-Caissier, 2016; Rana et al., 2020; Urain et al., 2020). In contrast, Pérez-Dattari and Kober (2023); Pérez-Dattari et al. (2024) enforce stability via additional loss functions derived using tools from the deep metric learning literature (Kaya and Bilge, 2019). Importantly, Pérez-Dattari et al. (2024) extends the ideas of Pérez-Dattari and Kober (2023) to more general scenarios, achieving better results in non-Euclidean state spaces and 2nd-order dynamical systems. These two features are essential for integrating such frameworks with fabrics, which, to the best of our knowledge, have not been simultaneously demonstrated in other works using time-invariant formulations, e.g. Rana et al. (2020); Mohammadi et al.; Saveriano et al. (2023).

From IL to whole-body motion generation For the safe operation of robots around humans, not only the dynamical system acting as the navigation policy should be stable but the whole-body motion is desired to be stable, collision-free, and within the system limits. However, most existing IL solutions that incorporate obstacle avoidance when learning dynamical systems in end-effector space focus solely on avoiding collisions with the end-effector, e.g. Mohammadi et al.; Ginesi et al. (2019); Huber et al. (2023). One possibility to achieve full-body collision avoidance in such cases is by employing collision-aware Inverse Kinematics (IK) solvers (Tedrake, 2023; Rakita et al., 2021).

Nevertheless, this ignores the desired acceleration or velocity profile specified by the learned dynamical system.

Another strategy is to combine IL with an optimization-based motion planner such as Model Predictive Control (MPC), inheriting the possibility to include kinodynamic and safety constraints and including the learned policy in the cost function. Several combinations exist, such as augmenting the cost function with a learned policy (Xiao et al., 2022), learning the complete cost function (Tagliabue and How, 2023), and learning the complete MPC framework (Ahn et al., 2023). Hu et al. (2023) propose MPDMP⁺ as a fast alternative for navigation amongst dynamic obstacles, combining DMPs with MPC where the cost function merges multiple demonstrations and includes obstacle avoidance using potential functions. Although some IL-MPC frameworks can guarantee stability (Ahn et al., 2023; Hewing et al., 2020), where most methods lack these guarantees, their applications remain limited with no demonstrations in real-world robotics.

Geometric motion generation A fast alternative strategy to optimization-based motion planners lies in the field of geometric motion generation. These techniques are inspired by Operational Space Control (Khatib, 1987), achieving stable and converging behavior for kinematically redundant robots using differential geometry (Bullo and Lewis, 2019). Recently, Riemannian Motion Policies and its extension, Geometric Fabrics, have been introduced (Ratliff et al., 2018; Ratliff and Van Wyk, 2023). In Geometric Fabrics, convergence is guaranteed with simple construction rules (Ratliff et al., 2020; Xie et al., 2020) and is also extended to dynamic environments (Spahn et al., 2023). These policies are shown to be beneficial in designing human-like motions (Klein et al., 2022) and have a high planning frequency compared to optimization-based methods (Spahn et al., 2023).

Combining geometric motion generation and learning Geometric methods have been considered in conjunction with reinforcement learning (Van Wyk et al., 2023) and Bayesian learning (Jaquier et al., 2020). Most interestingly for the present work, Xie et al. (2021, 2023) have looked into combining fabrics with IL. The method they propose is essentially different from TamedPUMA, as they directly learn the fabric, which results in limitations in terms of motion expressiveness.

2. Preliminaries

In this section, fundamental concepts are introduced for trajectory generation using artificial dynamical systems. Firstly, task and configuration spaces and their relations are specified in Section 2.1, followed by an overview of motion generation using geometric fabrics in Section 2.2. For a detailed overview of geometric fabrics and differential geometries, we refer the reader to Ratliff et al. (2020); Ratliff and Van Wyk (2023). Section 2.3 provides the necessary background on learned stable motion primitives and its stability properties.

2.1. Configuration and task spaces

The configuration of the robot is denoted by $\mathbf{q} \in \mathcal{C}$ with its time derivatives $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$. Here, \mathcal{C} indicates the configuration space of the robot, which has a dimension of n . Tasks can be defined in different task spaces \mathcal{X}_j . For instance, collision avoidance of a manipulator’s end effector can be a task defined in the end-effector’s space. Additionally, to address whole-body obstacle avoidance, multiple tasks can be defined for various coordinates along the robot’s body. Furthermore, other tasks, such as reaching behaviors, can also be included. A task variable $x_j \in \mathcal{X}_j$ denotes the value of the state representation for the j -th task space, where $j \in [M]$, M denotes the number of task

spaces, and $[M] = \{j \in \mathbb{Z}^+ : j \leq M\}$. The relation between the configuration space and a given task space is stated via a twice-differential map $\phi_j : \mathcal{C} \rightarrow \mathcal{X}_j$, and the map's Jacobian as $\mathbf{J}_{\phi_j} = \frac{\partial \phi_j}{\partial \mathbf{q}}$.

2.2. Geometric fabrics

A dynamical system describes the behavior of a system using differential equations. In *geometric fabrics* (Ratliff et al., 2020), these dynamical systems describe an *artificial* system generating desired trajectories for a robotic system. The desired motions are described using second-order nonlinear time-invariant dynamical systems, $\ddot{x} = f(\mathbf{x}, \dot{\mathbf{x}})$. In this framework, dynamical systems consist of two parts: one that conserves a *Finsler energy* and another that is energy-decreasing. The energy-conservative part takes care of all *avoidance tasks*, e.g., joint limit avoidance and obstacle avoidance, while the energy-decreasing part drives the system towards the goal.

Let us first focus on the energy-conservative dynamical system. The behavior for an avoidance task is described using a second-order dynamical system, $\mathbf{M}(\mathbf{x}, \dot{\mathbf{x}})\ddot{\mathbf{x}} + \boldsymbol{\xi}(\mathbf{x}, \dot{\mathbf{x}}) = \mathbf{0}$ within a task space \mathcal{X} where $\mathbf{M}(\mathbf{x}, \dot{\mathbf{x}})$ is symmetric and invertible. If this system, often referred to as a spectral semi-spray $\mathbf{h} : \mathcal{S} = (\mathbf{M}, \boldsymbol{\xi})_{\mathcal{X}}$, is conserving a Finsler energy, it is called a fabric $\tilde{\mathbf{h}} : \mathcal{S} = (\tilde{\mathbf{M}}, \tilde{\boldsymbol{\xi}})_{\mathcal{X}}$ (Ratliff et al., 2020). The system \mathbf{h} is transformed into a fabric $\tilde{\mathbf{h}}$ through *energization* with the Finsler energy, i.e., $\tilde{\mathbf{h}} = \text{energize}_{\mathcal{L}_e}[\mathbf{h}(\mathbf{x}, \dot{\mathbf{x}})]$ (Eq. 13 in Ratliff and Van Wyk (2023)). Notably, the system's energization only changes the speed along the path, but not the path itself, if the system is homogeneous of Degree 2, i.e., $\mathbf{h}(\mathbf{x}, \alpha \dot{\mathbf{x}}) = \alpha^2 \mathbf{h}(\mathbf{x}, \dot{\mathbf{x}})$, $\forall \alpha \geq 0$.

Each avoidance task can be described by these energy-conservative fabrics, $\ddot{\mathbf{x}}_j = \tilde{\mathbf{h}}(\mathbf{x}_j, \dot{\mathbf{x}}_j)$, within their own task space \mathcal{X}_j . To construct a whole-body policy for the robotic system, multiple task-space fabrics are mapped to the configuration space using a *pullback operation* and summed. The pullback operation, $\text{pull}_{\phi_j} \left(\tilde{\mathbf{M}}_j, \tilde{\boldsymbol{\xi}}_j \right)_{\mathcal{X}_j} = \left(\mathbf{J}_{\phi_j}^T \tilde{\mathbf{M}}_j \mathbf{J}_{\phi_j}, \mathbf{J}_{\phi_j}^T (\tilde{\boldsymbol{\xi}}_j + \dot{\mathbf{J}}_{\phi_j} \dot{\mathbf{q}}) \right)_{\mathcal{C}}$, is constructed as a function of ϕ_j , and maps the energy-conserving fabric to the configuration space, resulting in the system $\ddot{\mathbf{q}} = \tilde{\mathbf{h}}(\mathbf{q}, \dot{\mathbf{q}}) : \mathcal{S}_{\mathcal{C}} = (\tilde{\mathbf{M}}_j, \tilde{\boldsymbol{\xi}}_j)_{\mathcal{C}}$. The *specs* in \mathcal{C} are summed, e.g. $(\mathcal{S}_1 + \mathcal{S}_2)_{\mathcal{C}} = (\tilde{\mathbf{M}}_1 + \tilde{\mathbf{M}}_2, \tilde{\boldsymbol{\xi}}_1 + \tilde{\boldsymbol{\xi}}_2)_{\mathcal{C}}$ where the resulting dynamical system $\ddot{\mathbf{q}} = \tilde{\mathbf{h}}(\mathbf{q}, \dot{\mathbf{q}})$ is a fabric as well, since the summation and pullback operations are closed under algebra. This combined fabric can be forced towards the minimum of a potential $\psi(\mathbf{q}) \in \mathbb{R}$ and damped with a positive definite damping matrix (Ratliff et al., 2020). On a more abstract level, the desired dynamical system can be denoted as a combination of energy-conserving fabric $\tilde{\mathbf{h}}$ and an energy-decreasing navigation policy \mathbf{f} which is damped,

$$\ddot{\mathbf{q}} = \tilde{\mathbf{h}}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}). \quad (1)$$

2.3. Learning stable motion primitives via PUMA

Similarly to geometric fabrics, Policy via neUral Metric leArning (PUMA) models a desired motion as a nonlinear time-invariant dynamical system. This method represents the dynamical system $\mathbf{f}_{\theta}^{\mathcal{T}}$ in one of the task spaces \mathcal{X}_j (commonly the robot's end effector's space), denoted as \mathcal{T} , as a DNN with weights θ . The weights are optimized to imitate a set of demonstrations while ensuring convergence to a goal state $\mathbf{x}_g \in \mathcal{T}$ (Pérez-Dattari et al., 2024; Pérez-Dattari and Kober, 2023). In other words, \mathbf{x}_g must be a globally asymptotically stable equilibrium in the region of interest.

To enforce stability under minimization of the loss function, a specialized loss is introduced and optimized alongside an imitation loss. To design this loss, it is necessary first to define a latent space \mathcal{L} as the output of a hidden layer l of the DNN, such that

$$\ddot{\mathbf{x}} = \mathbf{f}_{\theta}^{\mathcal{T}}(\mathbf{x}, \dot{\mathbf{x}}) = \varphi_{\theta}(\rho_{\theta}(\mathbf{x}, \dot{\mathbf{x}})), \quad (2)$$

where $\rho_\theta : \mathcal{T} \rightarrow \mathcal{L}$ encodes the first $1, \dots, l$ layers and $\varphi_\theta : \mathcal{L} \rightarrow d\mathcal{T}$ the last $l+1, \dots, L$ layers with $d\mathcal{T}$ representing the tangent bundle of \mathcal{T} , capturing the derivatives and tangent spaces associated with \mathcal{T} . Then, a *mapped system* $f_\theta^{\mathcal{T} \rightarrow \mathcal{L}}(\mathbf{x}, \dot{\mathbf{x}})$ can be defined in \mathcal{L} as

$$\ddot{\mathbf{y}} = f_\theta^{\mathcal{T} \rightarrow \mathcal{L}}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{\partial \rho_\theta(\mathbf{x}, \dot{\mathbf{x}})}{\partial t}. \quad (3)$$

These definitions enable the formulation of the *stability conditions* introduced in Pérez-Dattari and Kober (2023), which can be stated as:

Theorem 1 (Stability conditions (Pérez-Dattari et al., 2024)) *In the region \mathcal{T} , \mathbf{x}_g is a globally asymptotically stable equilibrium of $f_\theta^{\mathcal{T}}$ if, $\forall \mathbf{x} \in \mathcal{T}$, (1) $\mathbf{y}_g = \rho_\theta(\mathbf{x}_g)$ is a globally asymptotically stable equilibrium of $f_\theta^{\mathcal{T} \rightarrow \mathcal{L}}$, and (2) $\rho_\theta(\mathbf{x}) = \mathbf{y}_g \Rightarrow \mathbf{x} = \mathbf{x}_g$.*

These conditions can be enforced via the following loss (Pérez-Dattari et al., 2024):

$$\ell_{\text{stable}} = \sum_{\mathbf{y}_0 \in \mathcal{B}} \sum_{t \in \mathcal{H}} \max(0, m + d(\mathbf{y}_g, \mathbf{y}(\mathbf{y}_0, t + \Delta t)) - d(\mathbf{y}_g, \mathbf{y}(\mathbf{y}_0, t))), \quad (4)$$

where $d(\cdot, \cdot)$ is a distance function, m is a small margin hyperparameter, and $\mathbf{y}(\mathbf{y}_0, t)$ represents the value of \mathbf{y} for an initial condition \mathbf{y}_0 and time t . Here, \mathcal{B} is a batch of initial conditions obtained at every iteration by randomly sampling the workspace \mathcal{T} , and \mathcal{H} contains different values of t up to some time horizon T . To imitate the set of demonstrations, this method minimized the combined loss $\ell_{\text{PUMA}} = \ell_{\text{IL}} + \lambda \ell_{\text{stable}}$, where ℓ_{IL} is the behavioral cloning loss employed in Pérez-Dattari and Kober (2023) and λ is a weight factor.

3. TamedPUMA: Combining learned stable motion primitives and fabrics

With learned stable motion primitives, complex tasks can be learned from demonstrations, while converging to the goal. In TamedPUMA, these learned dynamical systems are incorporated into the framework of geometric fabrics, generating stable and safe motions while respecting whole-body collision avoidance and physical constraints of the robot. In Section 3.1, the problem formulation is discussed, followed by the two variations of TamedPUMA, the Forcing Policy Method (FPM) and Compatible Potential Method (CPM), in Section 3.2 and 3.3 respectively.

3.1. Problem formulation

Consider there exists a dynamical system $\tilde{\mathbf{f}}$ that describes a desired motion profile for a robot, which induces the distribution of trajectories $\tilde{p}(\tau)$. For a given task variable $\mathbf{x} \in \mathcal{T}$, such as the robot's end-effector pose, we assume we can sample trajectories from this distribution. Then, if the dynamical system $f_\theta^{\mathcal{T}}$ represents the evolution of the robot's state in \mathcal{T} , our objective is to minimize the distance between the distribution of trajectories induced by this system $p_\theta(\tau)$ and $\tilde{p}(\tau)$. Furthermore, $f_\theta^{\mathcal{T}}$ must always reach a desired state $\mathbf{x}_g \in \mathcal{T}$ while accounting for region avoidance constraints, namely, self-collisions, obstacles, and joint limits for M tasks defined in multiple task spaces \mathcal{X}_j . More formally,

$$\mathbf{f}_\theta^{\mathcal{T}*} = \arg \min_{f_\theta^{\mathcal{T}}} D_{\text{KL}} (\tilde{p}(\tau) || p_\theta(\tau)), \quad (5a)$$

$$\text{s.t. } \lim_{t \rightarrow \infty} \|\mathbf{x} - \mathbf{x}_g\| = 0, \quad \forall \mathbf{x} \in \mathcal{T}, \quad (5b)$$

$$\mathbf{x}_j \in \mathcal{X}_j^{\text{free}}, \quad \forall j \in [M]. \quad (5c)$$

This objective minimizes the Kullback-Leibler divergence between the estimated dynamical system and the target system (5a). This divergence is often used as a distance measure between distributions as it is minimized through maximum likelihood estimation using distribution samples (Osa et al., 2018). Moreover, the problem is subject to two constraints: (5b) the dynamical system should converge to a task-space goal \mathbf{x}_g in \mathcal{T} , and (5c) task-space states $\mathbf{x}_j, \forall j \in [M]$, should remain within the free space $\mathcal{X}_j^{\text{free}}$ where the dynamical system is well-defined, which could be the collision-free space or the space within the joint limits.

From this formulation, PUMA softly addresses (5a) using the loss ℓ_{IL} , while fabrics addresses (5c) through their geometric-aware formulation. Both approaches tackle the problem of stability; however, it remains challenging to combine both methods while ensuring convergence to the goal (5b). In the following subsections, we propose two approaches to address this.

3.2. The Forcing Policy Method (FPM)

First, we introduce the FPM. For this purpose, we define the dynamical system \mathbf{f}_{θ}^C in configuration space resulting from applying a pullback operation, via PUMA, in \mathcal{T} :

$$\ddot{\mathbf{q}}_{\text{DNN}} = \mathbf{f}_{\theta}^C(\mathbf{q}, \dot{\mathbf{q}}) = \text{pull}_{\phi_{\mathcal{T}}} (\mathbf{f}_{\theta}^T(\mathbf{x}_{\text{ee}}, \dot{\mathbf{x}}_{\text{ee}})) , \quad (6)$$

where the task variable \mathbf{x}_{ee} is the end-effector position and orientation of the robot as illustrated in Fig. 1. Then, leveraging the definition of a forced system from Eq. (1), in the FPM we propose to use the pulled system obtained via PUMA as the forcing policy,

$$\ddot{\mathbf{q}}^d = \tilde{\mathbf{h}}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{f}_{\theta}^C(\mathbf{q}, \dot{\mathbf{q}}) . \quad (7)$$

Assuming ℓ_{PUMA} has already been minimized, the system \mathbf{f}_{θ}^T comes to rest at \mathbf{x}_g , implying that \mathbf{f}_{θ}^C converges to \mathbf{q}_g where multiple values of \mathbf{q}_g may exist in the case of a redundant system. This collection of states \mathbf{q}_g corresponds to the *zero set* of \mathbf{f}_{θ}^C . From Proposition II.17 in Ratliff and Van Wyk (2023), we know that if the system in Eq. (7) reaches the zero set of \mathbf{f}_{θ}^C , it will stay there (which comes from the observation that fabrics are conservative). However, convergence of Eq. (7) to the zero set of \mathbf{f}_{θ}^C is not formally assessed. Re-evaluating the problem formulation suggests that the constraint in Eq (5b), stating that the system should converge to the desired minimum \mathbf{q}_g , is therefore not guaranteed. In Sec. 3.3, we propose a method with stronger convergence guarantees.

3.3. The Compatible Potential Method (CPM)

As a second approach, we propose the CPM leveraging *compatible potentials* to obtain a stronger notion of convergence. For a potential function compatible with a dynamical system, its negative gradient generally points in the same direction as the system's vector field. More formally:

Definition 2 (Compatible potential (Ratliff and Van Wyk, 2023)) A potential function ψ is compatible with \mathbf{f} if: (1) $\partial\psi(\mathbf{q}) = \mathbf{0}$ if and only if $\mathbf{f}(\mathbf{q}, \mathbf{0}) = \mathbf{0}$, and (2) $-\partial\psi^\top \mathbf{f}(\mathbf{q}, \mathbf{0}) > \mathbf{0}$ wherever $\mathbf{f}(\mathbf{q}, \mathbf{0}) \neq \mathbf{0}$.

Building upon this, Theorem III.5 in Ratliff and Van Wyk (2023) states that given a dynamical system \mathbf{f} with a compatible potential, it is possible to construct another system that converges to the

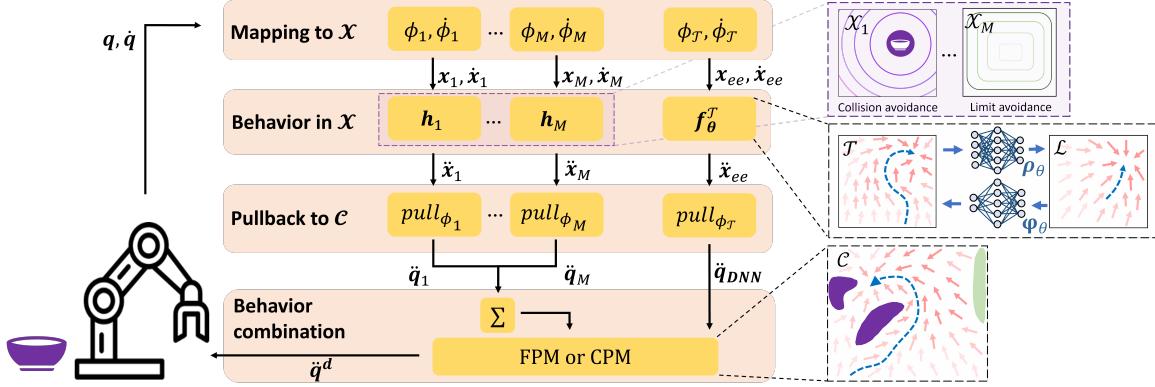


Figure 1: This illustration of TamedPUMA shows the behavior design given the relationships between the different task and configuration-space variables. The joint angles and velocities get mapped into task space where the desired behavior is specified. Via fabrics, all avoidance behaviors are defined using the joint limits, and varying obstacle positions, e.g. the position of the bowl. The DNN captures the desired behavior of the end-effector position and orientation.

zero set of f . Crucially, this new system can incorporate the avoidance behaviors (e.g., joint limits and collision avoidance) in configuration space, represented by $\mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) = \sum_{j=1}^M \mathbf{h}_j(\mathbf{q}, \dot{\mathbf{q}})$, as:

$$\ddot{\mathbf{q}}^d = \text{energize}_{\mathcal{H}}[\mathbf{h} + f] + \gamma(\mathbf{q}, \dot{\mathbf{q}}) \quad \text{with} \quad \gamma(\mathbf{q}, \dot{\mathbf{q}}) = - \left(\frac{\dot{\mathbf{q}} \dot{\mathbf{q}}^\top}{\dot{\mathbf{q}}^\top M_{\mathcal{L}_e} \dot{\mathbf{q}}} \right) \partial \psi - \beta \dot{\mathbf{q}}. \quad (8)$$

Here, we assume that the metric $M_{\mathcal{L}_e}$ is bounded in a finite region, strictly positive definite everywhere and does not vanish or reduce rank if $\dot{\mathbf{q}} \rightarrow \mathbf{0}$, $\beta > 0$, and obstacle positions and joint limits are static. Consequently, we aim to leverage this result by using $f_\theta^{\mathcal{C}}$ as the system f in Eq. (8) with the compatible potential. From the previous section, we already defined that the zero set of this system maps to the equilibrium $\mathbf{x}_g \in \mathcal{T}$. Hence, it remains to find a compatible potential for this function to employ the result from Theorem III.5.

Notably, if (4) is successfully minimized, it is possible to design a compatible potential for the system $f_\theta^{\mathcal{T}}$ in the latent space \mathcal{L} using the mapping ρ_θ by setting $\dot{\mathbf{x}} = \mathbf{0}$. Specifically, we can construct a potential using the latent-space variable \mathbf{y} and the encoder ρ_θ of PUMA as in Eq. (2),

$$\psi(\mathbf{x}) = \|\mathbf{y}_g - \mathbf{y}\|^2 = \|\rho_\theta(\mathbf{x}_g, \mathbf{0}) - \rho_\theta(\mathbf{x}, \mathbf{0})\|^2. \quad (9)$$

To observe that this is a compatible potential of $f_\theta^{\mathcal{T}}$, first, we highlight that since \mathbf{x}_g is asymptotically stable, we have $\partial \psi(\mathbf{x}_g) = \mathbf{0}$ if and only if $f(\mathbf{x}_g, \mathbf{0}) = \mathbf{0}$. This satisfies the first condition of Definition 2. Second, we note that for all $\mathbf{x} \neq \mathbf{x}_g$, Equation (4) ensures that the value of $\psi(\mathbf{x})$ decreases as $f_\theta^{\mathcal{T}}$ evolves over time, except for some edge cases that could, in theory, occur if ρ_θ has a large Lipschitz constant. These edge cases can be eliminated through regularization. Thus, this potential also satisfies the second condition of Definition 2 and therefore $\psi(\mathbf{x})$ is a compatible potential of $f_\theta^{\mathcal{T}}$. Finally, for Eq. 8 it only remains to express the gradient of this potential in configuration space, previously denoted as $\partial \psi$. For clarity, we will henceforth write this as $\partial \psi / \partial \mathbf{q}$. To achieve this, we require the forward kinematics from configuration space \mathcal{C} to task space \mathcal{T} , denoted $\phi^{\mathcal{T}}$. Then, we obtain

$$\partial \psi = \frac{\partial \psi}{\partial \mathbf{q}} = \frac{\partial \psi}{\partial \mathbf{x}} \cdot J_{\phi^{\mathcal{T}}}(\mathbf{x}), \quad (10)$$

where $J_{\phi}\tau$ is the Jacobian matrix of the forward kinematics, commonly available in robotic frameworks. The term $\partial\psi/\partial\mathbf{x}$ can be approximated via automatic differentiation tools for DNNs.

Assumptions on $M_{\mathcal{L}_e}$ To conclude this section, it is relevant to discuss further the implications of the assumptions on $M_{\mathcal{L}_e}$. For a fabric describing collision avoidance, two cases exist as the spec describing the fabric must be boundary conforming (Ratliff et al., 2020). (1) The metric $M_{\mathcal{L}_e}$ is finite along the eigen-directions parallel to the boundary’s tangent space but goes to infinity along directions orthogonal to the tangent space. (2) The metric $M_{\mathcal{L}_e}$ is a finite matrix along all trajectories, implying that $M_{\mathcal{L}_e}$ is also finite in the limit when $t \rightarrow \infty$. Considering our previous assumptions on $M_{\mathcal{L}_e}$ for (8), i.e. the metric is bounded in a finite region and strictly positive definite everywhere, we can design fabrics only according to the second scenario. This implies that, in the limit, we ensure convergence to the zero set of the forcing policy $f_\theta^C(\mathbf{q}, \dot{\mathbf{q}})$; however, collision avoidance is not guaranteed in the limit, as barrier-like functions that go to infinity on the boundary cannot be used to construct $M_{\mathcal{L}_e}$. For more details, we refer the reader to the attached material¹.

4. Experimental Results

In this section, we explore TamedPUMA’s capabilities in constructing stable and collision-free motions for a 7-DOF manipulator. Details regarding the DNN, differential mappings, a discussion on the limitations, code and videos are included in the attached material¹.

4.1. Experimental setup and performance metrics

To show the performance of the two variations of TamedPUMA, FPM and CPM, simulations using the Pybullet physics simulation (Coumans and Bai, 2016) and real-world experiments are performed on a 7-DoF KUKA iiwa manipulator. Two tasks are analyzed, picking a tomato from a crate and pouring liquid from a cup, where a DNN is trained for each task using 10 demonstrations, recording end-effector positions, velocities and accelerations. The proposed FPM and CPM, are compared against vanilla geometric fabrics, vanilla PUMA and *modulation-IK*. Modulation-IK modifies PUMA to be obstacle-free within the task space using a modulation matrix (Khansari-Zadeh and Billard, 2012). Then, whole-body collision avoidance is achieved by tracking the modified desired pose using a collision-aware IK (Tedrake, 2023). All methods are evaluated based on their *success rate* and *time-to-success*, which respectively indicate the ratio of successful scenarios and the time required for the robot to reach the goal pose. Successful scenarios ensure collision-free motions with an end pose satisfying $\|\mathbf{x}_{ee} - \mathbf{x}_g\|_2 < 0.05 \text{ m}$. In addition, the *computation time* is denoted, and the *path difference* to the desired path by PUMA as $\frac{1}{P} \sum \|\mathbf{x}_{ee} - \mathbf{x}_{PUMA}\|_2$ where \mathbf{x}_{ee} and \mathbf{x}_{PUMA} correspond to the end-effector poses along the path with length P of the analyzed method and PUMA respectively. The path difference is computed in obstacle-free scenarios, where we aim to track the DNN as closely as possible, and in obstacle-rich allowing for deviations from this path.

4.2. Simulation experiments on a 7-DOF manipulator

In simulation, 30 realistic scenarios are explored, including 15 scenarios of a tomato-picking task and 15 scenarios of a pouring task. In each task, the initial robot configuration and obstacle locations change. Moreover, 10 scenarios included moving goals, and 7 scenarios included moving obstacles. For each scenario, at least one obstacle is situated between the initial configuration and the goal

1. Link to the attached material: https://autonomousrobots.nl/paper_websites/pumafabrics

Table 1: Statistics for 30 simulated scenarios. The path difference to PUMA is measured in an obstacle-free environment, while all other metrics are compared in an obstacle-rich environment.

	Success-Rate	Time-to-Success [s]	Computation time [ms]	Path difference to PUMA
PUMA	0.40	4.75 ± 1.31	4.41 ± 0.26	0
Modulation-IK	0.57	5.75 ± 5.83	7.08 ± 5.98	0.39 ± 0.42
Fabrics	0.83	8.18 ± 5.54	0.39 ± 0.03	0.22 ± 0.27
FPM (ours)	1.00	6.97 ± 3.78	5.22 ± 0.22	0.02 ± 0.04
CPM (ours)	1.00	7.01 ± 3.57	6.55 ± 0.96	0.04 ± 0.06

pose. The last five links on the robotic chain are considered for collision avoidance, given that the first three links have restricted movement due to the base being mounted to the table. The shape of the last five links is approximated using spheres with a radius of 9 cm, and 7 cm for the final link.

As depicted in Table 1, the two TamedPUMA variations improve the success rate over PUMA by enabling whole-body obstacle avoidance. In contrast to geometric fabrics, FPM and CPM can track the desired motion profile leading to a smaller path difference with PUMA of 0.02 ± 0.04 and 0.04 ± 0.06 respectively, compared to geometric fabrics, 0.22 ± 0.27 , in an obstacle-free environment. In an obstacle-rich environment, geometric fabrics result in a deadlock in 6 of the 30 scenarios where the robot does not reach the goal as it is unable to move around the edge of the crate or object. The benchmark Modulation-IK is also unable to achieve all tasks due to collisions or deadlocks, as it cannot find a feasible solution converging towards the goal. TamedPUMA thereby inherits the efficient scalability to multi-object environments from fabrics, as computation time in an environment with 1000 obstacles is only 5.9 ± 0.8 ms and 7.0 ± 1.1 ms for FPM and CPM respectively, a neglectable increase with respect to the two obstacles considered in Table 1, while optimization-based methods like Modulation-IK scale poorly with large numbers of obstacles with a computation time of $3.5 \cdot 10^3 \pm 10.7 \cdot 10^3$ ms in an environment with 1000 obstacles. Although CPM offers stronger theoretical guarantees than FPM, performance is similar (see Table 1). Even though we do not optimize over the time-to-success, both FPM and CPM achieve the task within a reasonable time while remaining collision-free. Computation times are 4-7 ms on a standard laptop (i7-12700H) making the methodologies well suitable for real-time reactive motion generation in dynamic environments.

4.3. Real-world experiments on a 7-DOF manipulator

Experiments are performed on the real 7-DOF for the tomato-picking and pouring task where all obstacles, e.g. a bowl, a person’s hand and a helmet, are dynamically tracked in real-time via an optitrack system and modeled as spheres. In addition to the collision spheres considered during simulation, a collision sphere is added to the collision geometry on the center of the robotic hand with a diameter of 14 cm. The desired actions by TamedPUMA are sent at 30 Hz to a joint impedance controller running at 1000 Hz. Snapshots of a real-world experiment of the CPM are illustrated in Fig. 2, Fig. 3 and Fig 4, and the videos in the attached material show several movements for the simulated and real-world experiments using FPM and CPM. If the obstacles are not blocking the trajectory of the robot, the observed behavior of the proposed methods, FPM and CPM, are similar to PUMA and showcases clearly the learned behavior as demonstrated by the human. The user can push the robot away from the goal or change the goal online (Fig. 4), and TamedPUMA recovers from this disturbance. In the presence of obstacles, FPM and CPM achieve collision avoidance between the considered links on the robot and the obstacles while reaching the goal pose, as illustrated in Fig. 2, and Fig. 3.

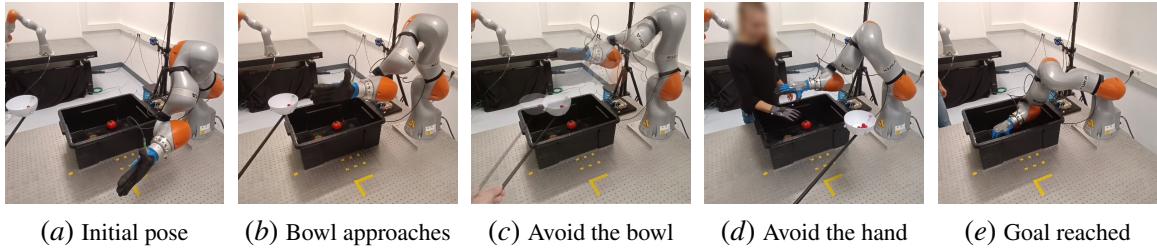


Figure 2: Selected time frames of CPM during a tomato-picking task with the bowl and hand as dynamic obstacles.

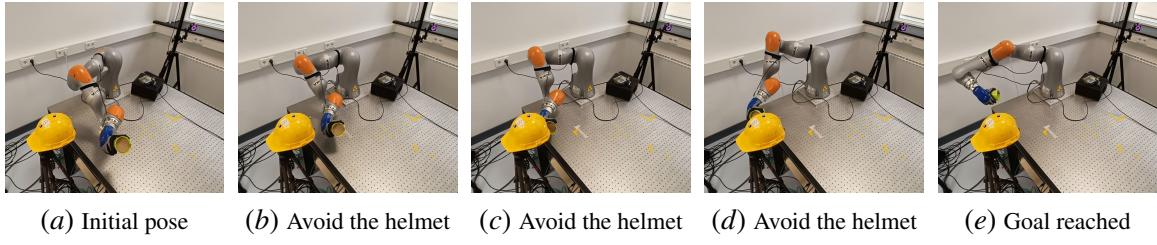


Figure 3: Selected time frames of CPM during a pouring task with the yellow helmet as a static obstacle.

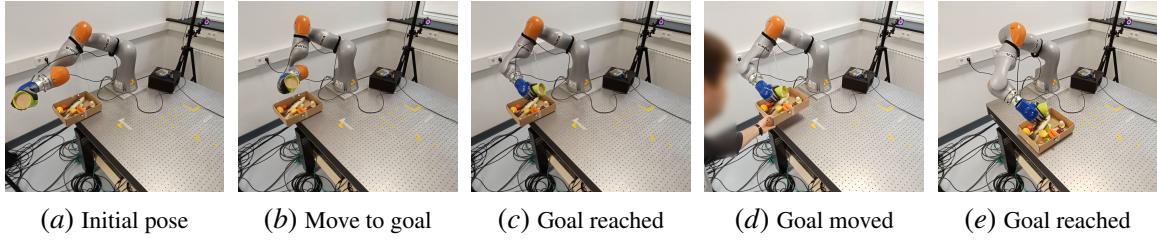


Figure 4: Selected time frames of CPM during a pouring task where the goal is changed online by the user.

A low-level joint impedance controller tracks the desired velocities and positions, outputting torque commands. This might decrease tracking performance compared to the simulated experiments, but allows users to push the robot away. If the user is no longer applying force to the robot, the system will converge to the goal pose, as illustrated in the videos in the attached material.

5. Conclusion

Imitation learning via stable motion primitives is a suitable approach for learning motion profiles from demonstrations while providing convergence to the goal. We introduced TamedPUMA, a safe and stable extension of learned stable motion primitives augmented with the recently developed geometric fabrics for safe and stable operations in the presence of obstacles. We proposed two variations, the Forcing Policy Method and Compatible Potential Method, ensuring respectively that the goal is stable, or the stronger notion that the system converges towards the reachable goal. Experiments were carried out both in simulation and in the real world. When trained on a tomato-picking task or pouring task, the proposed TamedPUMA generates a desired motion profile using a DNN while taking whole-body collision avoidance and joint limits into account, with a computation time of just 4-7 ms.

Acknowledgments

This project has received funding from the European Union through ERC, INTERACT, under Grant 101041863, and by the NXTGEN national program. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the NXTGEN national program. Neither the European Union nor the granting authority can be held responsible for them.

References

- Kwangjun Ahn, Zakaria Mhammedi, Horia Mania, Zhang-Wei Hong, and Ali Jadbabaie. Model predictive control via on-policy imitation learning. In *Learning for Dynamics and Control Conference*, pages 1493–1505. PMLR, 2023.
- David Bao, S-S Chern, and Zhongmin Shen. *An introduction to Riemann-Finsler geometry*, volume 200. Springer Science & Business Media, 2012.
- Aude Billard, Sina Mirrazavi, and Nadia Figueiroa. *Learning for adaptive and reactive robot control: a dynamical systems approach*. Mit Press, 2022.
- Francesco Bullo and Andrew D Lewis. *Geometric control of mechanical systems: modeling, analysis, and design for simple mechanical control systems*, volume 49. Springer, 2019.
- Erwin Coumans and Yunfei Bai. Pybullet, a python module for physics simulation for games, robotics and machine learning. <http://pybullet.org>, 2016.
- NB Figueiroa Fernandez and A Billard. A physically-consistent bayesian non-parametric mixture model for dynamical system learning. *Proceedings of Machine Learning Research*, 2018.
- Michele Ginesi, Daniele Meli, Andrea Calanca, Diego Dall’Alba, Nicola Sansonetto, and Paolo Fiorini. Dynamic movement primitives: Volumetric obstacle avoidance. In *2019 19th international conference on advanced robotics (ICAR)*, pages 234–239. IEEE, 2019.
- Lukas Hewing, Kim P Wabersich, Marcel Menner, and Melanie N Zeilinger. Learning-based model predictive control: Toward safe learning in control. *Annual Review of Control, Robotics, and Autonomous Systems*, 3:269–296, 2020.
- Yingbai Hu, Mingyang Cui, Jianghua Duan, Wenjun Liu, Dianye Huang, Alois Knoll, and Guang Chen. Model predictive optimization for imitation learning from demonstrations. *Robotics and Autonomous Systems*, 163:104381, 2023.
- Yingbai Hu, Fares J Abu-Dakka, Fei Chen, Xiao Luo, Zheng Li, Alois Knoll, and Weiping Ding. Fusion dynamical systems with machine learning in imitation learning: A comprehensive overview. *Information Fusion*, page 102379, 2024.
- Lukas Huber, Jean-Jacques Slotine, and Aude Billard. Avoidance of concave obstacles through rotation of nonlinear dynamics. *IEEE Transactions on Robotics*, 2023.

- Auke Jan Ijspeert, Jun Nakanishi, Heiko Hoffmann, Peter Pastor, and Stefan Schaal. Dynamical movement primitives: learning attractor models for motor behaviors. *Neural computation*, 25(2):328–373, 2013.
- Noémie Jaquier, Leonel Rozo, Sylvain Calinon, and Mathias Bürger. Bayesian optimization meets riemannian manifolds in robot learning. In *Conference on Robot Learning*, pages 233–246. PMLR, 2020.
- Mahmut Kaya and Hasan Şakir Bilge. Deep metric learning: A survey. *Symmetry*, 11(9):1066, 2019.
- S Mohammad Khansari-Zadeh and Aude Billard. Learning stable nonlinear dynamical systems with gaussian mixture models. *IEEE Transactions on Robotics*, 27(5):943–957, 2011.
- Seyed Mohammad Khansari-Zadeh and Aude Billard. A dynamical system approach to realtime obstacle avoidance. *Autonomous Robots*, 32:433–454, 2012.
- Oussama Khatib. A unified approach for motion and force control of robot manipulators: The operational space formulation. *IEEE Journal on Robotics and Automation*, 3(1):43–53, 1987.
- Holger Klein, Noémie Jaquier, Andre Meixner, and Tamim Asfour. A riemannian take on human motion analysis and retargeting. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5210–5217. IEEE, 2022.
- Andre Lemme, Klaus Neumann, René Felix Reinhart, and Jochen J Steil. Neural learning of vector fields for encoding stable dynamical systems. *Neurocomputing*, 141:3–14, 2014.
- Ge Li, Zeqi Jin, Michael Volpp, Fabian Otto, Rudolf Lioutikov, and Gerhard Neumann. Prodmp: A unified perspective on dynamic and probabilistic movement primitives. *IEEE Robotics and Automation Letters*, 8(4):2325–2332, 2023.
- Hadi Beik Mohammadi, Søren Hauberg, Georgios Arvanitidis, Nadia Figueroa, Gerhard Neumann, and Leonel Rozo. Neural contractive dynamical systems. In *The Twelfth International Conference on Learning Representations*.
- Takayuki Osa, Joni Pajarinen, Gerhard Neumann, J Andrew Bagnell, Pieter Abbeel, Jan Peters, et al. An algorithmic perspective on imitation learning. *Foundations and Trends® in Robotics*, 7(1-2):1–179, 2018.
- Rodrigo Pérez-Dattari and Jens Kober. Stable motion primitives via imitation and contrastive learning. *IEEE Transactions on Robotics*, 39(5):3909–3928, 2023.
- Rodrigo Pérez-Dattari, Cosimo Della Santina, and Jens Kober. Puma: Deep metric imitation learning for stable motion primitives. *Advanced Intelligent Systems*, page 2400144, 2024.
- Nicolas Perrin and Philipp Schlehuber-Caissier. Fast diffeomorphic matching to learn globally asymptotically stable nonlinear dynamical systems. *Systems & Control Letters*, 96:51–59, 2016.
- Affan Pervez, Yuecheng Mao, and Dongheui Lee. Learning deep movement primitives using convolutional neural networks. In *2017 IEEE-RAS 17th international conference on humanoid robotics (Humanoids)*, pages 191–197. IEEE, 2017.

- Michael Przystupa, Faezeh Haghverd, Martin Jagersand, and Samuele Tosatto. Deep probabilistic movement primitives with a bayesian aggregator. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3704–3711. IEEE, 2023.
- Daniel Rakita, Haochen Shi, Bilge Mutlu, and Michael Gleicher. Collisionik: A per-instant pose optimization method for generating robot motions with environment collision avoidance. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9995–10001. IEEE, 2021.
- Muhammad Asif Rana, Anqi Li, Dieter Fox, Byron Boots, Fabio Ramos, and Nathan Ratliff. Euclideanizing flows: Diffeomorphic reduction for learning stable dynamical systems. In *Learning for Dynamics and Control*, pages 630–639. PMLR, 2020.
- Nathan Ratliff and Karl Van Wyk. Fabrics: A foundationally stable medium for encoding prior experience. *arXiv preprint arXiv:2309.07368*, 2023.
- Nathan D Ratliff, Jan Issac, Daniel Kappler, Stan Birchfield, and Dieter Fox. Riemannian motion policies. *arXiv preprint arXiv:1801.02854*, 2018.
- Nathan D Ratliff, Karl Van Wyk, Mandy Xie, Anqi Li, and Muhammad Asif Rana. Optimization fabrics. *arXiv preprint arXiv:2008.02399*, 2020.
- Barry Ridge, Andrej Gams, Jun Morimoto, Aleš Ude, et al. Training of deep neural networks for the generation of dynamic movement primitives. *Neural Networks*, 127:121–131, 2020.
- Matteo Saveriano, Fares J Abu-Dakka, and Ville Kyrki. Learning stable robotic skills on riemannian manifolds. *Robotics and Autonomous Systems*, 169:104510, 2023.
- Max Spahn, Martijn Wisse, and Javier Alonso-Mora. Dynamic optimization fabrics for motion generation. *IEEE Transactions on Robotics*, 2023.
- Andrea Tagliabue and Jonathan P How. Efficient deep learning of robust policies from mpc using imitation and tube-guided data augmentation. *arXiv preprint arXiv:2306.00286*, 2023.
- Russ Tedrake. *Robotic Manipulation*. 2023. URL <http://manipulation.mit.edu>.
- Aleš Ude, Bojan Nemeć, Tadej Petrić, and Jun Morimoto. Orientation in cartesian space dynamic movement primitives. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2997–3004. IEEE, 2014.
- J Urain, M Ginesi, D Tateo, and J Peters. Imitationflow: Learning deep stable stochastic dynamic systems by normalizing flows. In *RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5231–5237, 2020.
- Karl Van Wyk, Ankur Handa, Viktor Makoviychuk, Yujie Guo, Arthur Allshire, and Nathan D. Ratliff. Geometric fabrics: a safe guiding medium for policy learning. *arXiv preprint arXiv:2310.12831*, 2023.

Xuesu Xiao, Tingnan Zhang, Krzysztof Choromanski, Edward Lee, Anthony Francis, Jake Varley, Stephen Tu, Sumeet Singh, Peng Xu, Fei Xia, et al. Learning model predictive controllers with real-time attention for real-world navigation. *Proceedings of The 6th Conference on Robot Learning*, 205:1708-1721, 2022.

Mandy Xie, Karl Van Wyk, Anqi Li, Muhammad Asif Rana, Qian Wan, Dieter Fox, Byron Boots, and Nathan Ratliff. Geometric fabrics for the acceleration-based design of robotic motion. *arXiv preprint arXiv:2010.14750*, 2020.

Mandy Xie, Anqi Li, Karl Van Wyk, Frank Dellaert, Byron Boots, and Nathan Ratliff. Imitation learning via simultaneous optimization of policies and auxiliary trajectories. *arXiv preprint arXiv:2105.03019*, 2021.

Mandy Xie, Ankur Handa, Stephen Tyree, Dieter Fox, Harish Ravichandar, Nathan D Ratliff, and Karl Van Wyk. Neural geometric fabrics: Efficiently learning high-dimensional policies from demonstration. In *Conference on Robot Learning*, pages 1355–1367. PMLR, 2023.