Proceedings Paper:
Huang, Y, Silvério, J and Caldwell, DG (2019) Towards Minimal Intervention Control with Competing Constraints. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IROS 2018, 01-05 Oct 2018, Madrid, Spain. IEEE , pp. 733-738.

https://doi.org/10.1109/IROS.2018.8594235

Towards Minimal Intervention Control with Competing Constraints

Yanlong Huang, João Silvério and Darwin G. Caldwell

***Abstract*— As many imitation learning algorithms focus on pure trajectory generation in either Cartesian space or joint space, the problem of considering competing trajectory constraints from both spaces still presents several challenges. In particular, when perturbations are applied to the robot, the underlying controller should take into account the importance of each space for the task execution, and compute the control effort accordingly. However, no such controller formulation exists. In this paper, we provide a minimal intervention control strategy that simultaneously addresses the problems of optimal control and competing constraints between Cartesian and joint spaces. In light of the inconsistency between Cartesian and joint constraints, we exploit the robot null space from an information-theory perspective so as to reduce the corresponding conflict. An optimal solution to the aforementioned controller is derived and furthermore a connection to the classical finite horizon linear quadratic regulator (LQR) is provided. Finally, a writing task in a simulated robot verifies the effectiveness of our approach.**

I. Introduction

In the past few years, imitation learning has been studied in a myriad of applications, such as pouring tasks [1], striking motions [2] and obstacle avoidance [3]. While many approaches focus on skill learning in either Cartesian space or joint space, an important problem arises: can robots imitate human skills in both Cartesian and joint spaces simultaneously?

Indeed, learning in a single space has achieved remarkable performances in many systems. However, for tasks where the importance of Cartesian and joint trajectories varies with time, the learning in a single space might be inappropriate. To take a sequential bottle-shaking task as an example [4], it requires the robot to first reach and grasp a bottle, then shake it. During the reaching and grasping phase, the Cartesian trajectory is more important since the robot end-effector needs to reach the bottle precisely, and subsequently in the shaking phase the joint trajectory plays a crucial role since certain joints govern the shaking movement.

For the cases where Cartesian and joint trajectories are equally important, hybrid imitation learning in both spaces becomes highly desirable. A typical example is the striking movement in a robot table tennis scenario. As suggested in [5], the robot Cartesian trajectory (i.e., trajectory of the racket that is attached to the robot end-effector) should coincide with the ball trajectory but with an opposite direction, so that a higher interception rate can be achieved. Besides, for the sake of producing a natural striking movement, robot joint configuration (i.e., striking posture) also needs to be determined properly.

A few approaches have been proposed for the hybrid imitation learning, such as [4], [6], [7]. However, they only consider the imitation of position or velocity trajectories, becoming, in essence, a pure trajectory generation problem. In fact, from a control perspective, it would be desired to compute proper control commands so as to drive the robot to imitate demonstrated skills (in terms of position and velocity trajectories) in both Cartesian and joint spaces. In this paper, we aim for designing a controller in robot joint space that allows robots to mimic demonstrated skills in both Cartesian and joint spaces. Similarly to [6], we refer to demonstrations in both spaces as *competing constraints* throughout this paper, since the robot motion is required to follow both constraints.

Inspired by [8], [9], we propose to formulate the aforementioned issues into a minimal intervention control problem that incorporates competing constraints. However, due to the fact that Cartesian constraints and joint control commands are non-linearly dependent, it is non-trivial to design the optimal controller. A straightforward way is to transform Cartesian constraints into joint space, and subsequently, reformulate the competing constraints as two constraints in joint space, rendering the minimal intervention control problem feasible. In [4], [6], Jacobian-based inverse kinematics was used to map Cartesian constraints into joint space. However, the robot null space is ignored in both works, which might yield a new joint constraint that is inconsistent with the original joint constraint. Thus, we propose to exploit the robot null space towards reducing the possible conflict from both spaces.

In comparison to the aforementioned state-of-the-art approaches, our main contributions are:

(i) a minimal intervention control framework with competing constraints,
(ii) an information-theory perspective introduced to guide the optimization of the robot null space towards reducing the conflict between competing constraints,
(iii) a dual interpretation that connects the control problem with competing constraints and the finite horizon LQR.

This paper is arranged as follows. We first explain the estimation of probabilistic reference trajectories in Section II. Subsequently, we propose a minimal intervention controller which incorporates competing constraints described by reference trajectories and derive an approximated solution for this controller in Section III. A dual interpretation is provided, connecting our approach to the typical finite horizon LQR in Section IV. Evaluations of our framework in a seven

All authors are with Department of Advanced Robotics, Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genoa, Italy, firstname.lastname@iit.it

This work was supported by the Italian Ministry of Defense.

degree-of-freedom (DoF) simulated robot are presented in Section V. Finally, we conclude this paper and discuss the shortcomings and possible extensions of our work in Section VI.

II. Probabilistic Modeling of Demonstrations

Inspired by previous works [8], [9] that exploit the probabilistic properties underlying multiple demonstrations to design optimal controllers, we here propose to model the demonstrations from Cartesian and joint spaces probabilistically. Let us denote M demonstrations in Cartesian space and joint space as $\{\{t_{n,m}, \mathbf{p}_{n,m}, \dot{\mathbf{p}}_{n,m}\}_{n=1}^{N}\}_{m=1}^{M}$ and $\{\{t_{n,m}, \mathbf{q}_{n,m}, \dot{\mathbf{q}}_{n,m}\}_{n=1}^{N}\}_{m=1}^{M}$, respectively, where each demonstration has length N, $\mathbf{p}_{n,m} \in \mathbb{R}^3$ and $\dot{\mathbf{p}}_{n,m}$ respectively denote three-dimensional Cartesian position and velocity, $\mathbf{q}_{n,m} \in \mathbb{R}^{\mathcal{O}}$ and $\dot{\mathbf{q}}_{n,m}$ represent $\mathcal{O}$-dimensional vector of joint positions and velocities, respectively. Since the datapoints in both spaces are high-dimensional, we employ Gaussian mixture model (GMM) [10], [11] to model the joint probability distribution $\mathcal{P}(t, \mathbf{p}, \dot{\mathbf{p}})$ and $\mathcal{P}(t, \mathbf{q}, \dot{\mathbf{q}})$. Note that GMM is also applied in [4] and [6], where the former models positions while the latter models velocities.

In order to facilitate the explanation, we here take the modeling of Cartesian trajectories as an example, and address the joint space afterwards. Formally, we write $\boldsymbol{\xi}_t = [\mathbf{p}_t^T \ \dot{\mathbf{p}}_t^T]^T$ and $\mathcal{P}(t, \boldsymbol{\xi}_t)$ as

$$\mathcal{P}(t, \boldsymbol{\xi}_t) \sim \sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \tag{1}$$

where π_k denotes the prior probability of each Gaussian component, $\boldsymbol{\mu}_k = \begin{bmatrix} \boldsymbol{\mu}_{t,k} \\ \boldsymbol{\mu}_{\xi,k} \end{bmatrix}$ and $\boldsymbol{\Sigma}_k = \begin{bmatrix} \boldsymbol{\Sigma}_{tt,k} & \boldsymbol{\Sigma}_{t\xi,k} \\ \boldsymbol{\Sigma}_{\xi t,k} & \boldsymbol{\Sigma}_{\xi\xi,k} \end{bmatrix}$ represent mean and covariance of the i-th Gaussian component, respectively. With the model (1), we use Gaussian mixture regression (GMR) to retrieve probabilistic reference trajectories. Given a query point t, its corresponding trajectory point $\boldsymbol{\xi}(t)$ is computed as [10]

$$\boldsymbol{\xi}(t) \sim \sum_{k=1}^{K} h_k(t) \mathcal{N}\left(\boldsymbol{\mu}_k(t), \boldsymbol{\Sigma}_k\right) \tag{2}$$

with

$$h_k(t) = \frac{\pi_k \mathcal{P}(t|\boldsymbol{\mu}_{t,k}, \boldsymbol{\Sigma}_{tt,k})}{\sum_{i=1}^{K} \pi_i \mathcal{P}(t|\boldsymbol{\mu}_{t,i}, \boldsymbol{\Sigma}_{tt,i})}, \tag{3}$$

$$\boldsymbol{\mu}_k(t) = \boldsymbol{\mu}_{\xi,k} + \boldsymbol{\Sigma}_{\xi t,k} \boldsymbol{\Sigma}_{tt,k}^{-1} (t - \boldsymbol{\mu}_{t,k}), \tag{4}$$

$$\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}_{\xi\xi,k} - \boldsymbol{\Sigma}_{\xi t,k} \boldsymbol{\Sigma}_{tt,k}^{-1} \boldsymbol{\Sigma}_{t\xi,k}. \tag{5}$$

Note that (2) can be approximated by a single Gaussian distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_t^{\xi}, \hat{\boldsymbol{\Sigma}}_t^{\xi})$ (refer to [10] for details). For the sake of convenient discussion, we write the retrieved reference datapoint in Cartesian space at time t as $\hat{\boldsymbol{\xi}}_t$, which satisfies $\mathcal{P}(\hat{\boldsymbol{\xi}}_t|t) \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_t^{\xi}, \hat{\boldsymbol{\Sigma}}_t^{\xi})$.

The modeling of demonstrations in joint space can be carried out in a similar way. In this case, we write $\boldsymbol{\zeta}_t = [\mathbf{q}_t^T \ \dot{\mathbf{q}}_t^T]^T$ and the corresponding reference datapoint in joint space at time t as $\hat{\boldsymbol{\zeta}}_t$ with $\mathcal{P}(\hat{\boldsymbol{\zeta}}_t|t) \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_t^{\zeta}, \hat{\boldsymbol{\Sigma}}_t^{\zeta})$. In the next section, we exploit the probabilistic reference trajectories in Cartesian and joint spaces to design a controller, that takes into account competing constraints.

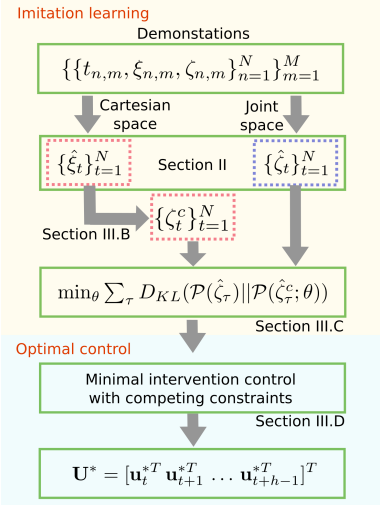

Fig. 1. An overview of the minimal intervention control with competing constraints. Given demonstrated trajectories, GMM/GMR are employed to generate the corresponding reference trajectories in Cartesian and joint spaces, respectively. Then, the robot null space is exploited to reduce the conflict between the transformed and demonstrated joint distributions. Finally, the proposed control problem is solved similarly to the finite horizon LQR.

III. Minimal Intervention Control with Competing Constraints

Since the probabilistic reference trajectories $\{\hat{\boldsymbol{\xi}}_t\}_{t=1}^{N}$ and $\{\hat{\boldsymbol{\zeta}}_t\}_{t=1}^{N}$ encapsulate the distributions of demonstrated trajectories in Cartesian and joint spaces, we aim at designing a controller which can incorporate both reference trajectories (i.e., competing constraints) while employing control efforts with optimal amplitude, given the task constraints in both spaces. Formally, we formulate this issue as a minimal intervention control problem with competing constraints (Section III-A). Subsequently, we exploit the robot null space to reduce the inconsistency between both reference trajectories (Section III-B and III-C). Finally, we provide an approximated solution to the proposed controller (Section III-D). Our framework is illustrated in Fig. 1.

A. Problem Description

Since the robot is controlled in its joint space, we here study the problem of finding proper joint control commands $\mathbf{u}$ such that the competing constraints are fulfilled. Specifically, we formulate this problem as minimizing

$$J(\mathbf{U}) = \sum_{\tau=t}^{t+h} (\boldsymbol{\xi}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\xi})^T \mathbf{Q}_\tau^{\xi} (\boldsymbol{\xi}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\xi}) + (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\zeta})^T \mathbf{Q}_\tau^{\zeta} (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\zeta}) + \sum_{\tau=t}^{t+h-1} \mathbf{u}_\tau^T \mathbf{R}_\tau \mathbf{u}_\tau, \tag{6}$$

where $\mathbf{U} = [\mathbf{u}_t^T \ \mathbf{u}_{t+1}^T \ \dots \ \mathbf{u}_{t+h-1}^T]^T$ represents the sequence of control commands and $h > 0$ denotes the horizon length

of prediction. The weight matrices $\mathbf{Q}_\tau^\xi$, $\mathbf{Q}_\tau^\zeta$ and $\mathbf{R}_\tau$ are positive-definite. Note that robot joint acceleration is related closely to the trajectory smoothness and torque limit, thus we consider the control command as $\mathbf{u} = \ddot{\mathbf{q}}$. Similarly to [8], [9], we exploit the trajectory variations in Cartesian and joint spaces to design a *minimal intervention controller*. Namely, we set $\mathbf{Q}_\tau^\xi = (\hat{\mathbf{\Sigma}}_\tau^\xi)^{-1}$ and $\mathbf{Q}_\tau^\zeta = (\hat{\mathbf{\Sigma}}_\tau^\zeta)^{-1}$. In this setting, the covariance of the probabilistic reference trajectory is viewed as an importance measure for the tracking problem, which implies that a large tracking error is allowed for the large covariance while a small error is required when the covariance is small.

Differing from the approaches in [4], [6] that only consider the position or velocity constraints in Cartesian space and joint spaces, we consider a control problem (as defined by (6)) which not only requires to imitate positions and velocities in both spaces, but also includes a penalty of the control commands. In contrast to [9] that designs a minimal intervention controller in either task space or joint space, the proposed control problem (6) is capable of incorporating competing constraints from both Cartesian and joint spaces and meanwhile preferring small control efforts at the control level.

B. Transformation of Cartesian Constraints

In comparison with the finite horizon LQR that typically relies on a linear dynamics model, it is non-trivial to solve (6) since a non-linear forward kinematics is involved. In order to make the problem in (6) tractable, we transform the Cartesian constraint into the joint space. Assuming that robot Jacobian function $\mathbf{J}(\mathbf{q})$ is available, we have

$$\begin{aligned} \mathbf{q}_t &= \mathbf{q}_{t-1} + \mathbf{J}^\dagger(\mathbf{p}_t - \mathbf{p}_{t-1}) + (\mathbf{I} - \mathbf{J}^\dagger\mathbf{J})\mathbf{M}(\boldsymbol{\theta})\delta_t \\ \dot{\mathbf{q}}_t &= \mathbf{J}^\dagger\dot{\mathbf{p}}_t + (\mathbf{I} - \mathbf{J}^\dagger\mathbf{J})\mathbf{M}(\boldsymbol{\theta}) \end{aligned}, \tag{7}$$

where $\mathbf{J}^\dagger = \mathbf{J}^T(\mathbf{J}\mathbf{J}^T)^{-1}$, $\delta_t > 0$ denotes the time interval and $\mathbf{M}(\boldsymbol{\theta}) = \mathbf{\Phi}(t)^T\boldsymbol{\theta}$ represents the null space with basis functions $\mathbf{\Phi}(t)$ and associated hyper-parameters $\boldsymbol{\theta}$. For simplicity, we write $\mathbf{J}$ instead of $\mathbf{J}(\mathbf{q}_{t-1})$. Furthermore, (7) can be rewritten in a compact form as

$$\underbrace{\begin{bmatrix}\mathbf{q}_t\\ \dot{\mathbf{q}}_t\end{bmatrix}}_{\hat{\boldsymbol{\zeta}}_t^c} = \underbrace{\begin{bmatrix}\mathbf{J}^\dagger & \mathbf{0}\\ \mathbf{0} & \mathbf{J}^\dagger\end{bmatrix}}_{\mathbf{J}_1}\underbrace{\begin{bmatrix}\mathbf{p}_t\\ \dot{\mathbf{p}}_t\end{bmatrix}}_{\hat{\boldsymbol{\xi}}_t} + \underbrace{\begin{bmatrix}\mathbf{I}-\mathbf{J}^\dagger\mathbf{J} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}-\mathbf{J}^\dagger\mathbf{J}\end{bmatrix}}_{\mathbf{J}_2}\underbrace{\begin{bmatrix}\mathbf{M}(\boldsymbol{\theta})\delta_t\\ \mathbf{M}(\boldsymbol{\theta})\end{bmatrix}}_{\hat{\mathbf{M}}(\boldsymbol{\theta})} + \underbrace{\begin{bmatrix}\mathbf{q}_{t-1} - \mathbf{J}^\dagger\mathbf{p}_{t-1}\\ \mathbf{0}\end{bmatrix}}_{\mathbf{C}}. \tag{8}$$

On the basis of the distribution of $\hat{\boldsymbol{\xi}}_t$, we can estimate the transformed joint constraint $\hat{\boldsymbol{\zeta}}_t^c$ whose expectation and covariance are

$$\begin{aligned} \hat{\boldsymbol{\mu}}_t^c &= \mathbb{E}(\hat{\boldsymbol{\zeta}}_t^c) = \mathbf{J}_1\hat{\boldsymbol{\mu}}_t^\xi + \mathbf{J}_2\hat{\mathbf{M}}(\boldsymbol{\theta}) + \mathbf{C} \\ \hat{\mathbf{\Sigma}}_t^c &= \mathbb{D}(\hat{\boldsymbol{\zeta}}_t^c) = \mathbf{J}_1\hat{\mathbf{\Sigma}}_t^\xi\mathbf{J}_1^T. \end{aligned} \tag{9}$$

It is noted that the new joint constraint depends on the robot null space parameters $\boldsymbol{\theta}$. Unlike previous work in [4], [6] that neglects the robot null space, we exploit $\mathbf{M}(\boldsymbol{\theta})$ to enrich the transformation from Cartesian space into joint space.

C. Optimization Criterion for Null-Space Parameters

Let us first revisit the problem formulation in (6). Since the Cartesian constraint $\hat{\boldsymbol{\xi}}_t$ is transformed into a new joint constraint $\hat{\boldsymbol{\zeta}}_t^c$, we consider the following objectives

$$\begin{aligned} e_1 &= \sum_{\tau=t}^{t+h}(\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^c)^T(\hat{\mathbf{\Sigma}}_\tau^c)^{-1}(\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^c) \quad \text{and} \\ e_2 &= \sum_{\tau=t}^{t+h}(\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^\zeta)^T(\hat{\mathbf{\Sigma}}_\tau^\zeta)^{-1}(\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^\zeta). \end{aligned} \tag{10}$$

With the definition of multivariate Gaussian distribution [12], the minimization of e_1 and e_2 are equivalent to the maximization of $m_1 = \prod_{\tau=t}^{t+h}\mathcal{N}(\boldsymbol{\zeta}_\tau|\hat{\boldsymbol{\mu}}_\tau^c, \hat{\mathbf{\Sigma}}_\tau^c)$ and $m_2 = \prod_{\tau=t}^{t+h}\mathcal{N}(\boldsymbol{\zeta}_\tau|\hat{\boldsymbol{\mu}}_\tau^\zeta, \hat{\mathbf{\Sigma}}_\tau^\zeta)$, respectively. However, due to the possible conflict between both maximum likelihood problems, it is almost impossible to maximize m_1 and m_2 simultaneously. Thus, an intuitive way to do it is to minimize the inconsistency between $\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^c, \hat{\mathbf{\Sigma}}_\tau^c)$ and $\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^\zeta, \hat{\mathbf{\Sigma}}_\tau^\zeta)$. As a natural approach to measure the distance between two probability distributions, the well-known *Kullback-Leibler* (KL) divergence [12], [13] has been applied in many areas such as policy search [14] and trajectory optimization [15]. Here, we minimize the KL-divergence based objective

$$J_{kl}(\boldsymbol{\theta}) = \sum_{\tau=1}^{N} D_{KL}\left(\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^\zeta, \hat{\mathbf{\Sigma}}_\tau^\zeta)||\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^c, \hat{\mathbf{\Sigma}}_\tau^c; \boldsymbol{\theta})\right) \tag{11}$$

with

$$\begin{aligned} D_{KL}\big(\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^\zeta, \hat{\mathbf{\Sigma}}_\tau^\zeta)||\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^c, \hat{\mathbf{\Sigma}}_\tau^c)\big) = \frac{1}{2}\Bigg(\log\frac{|\hat{\mathbf{\Sigma}}_\tau^c|}{|\hat{\mathbf{\Sigma}}_\tau^\zeta|} - \mathcal{O} \\ + \mathrm{Tr}\Big((\hat{\mathbf{\Sigma}}_\tau^c)^{-1}\hat{\mathbf{\Sigma}}_\tau^\zeta\Big) + (\hat{\boldsymbol{\mu}}_\tau^c - \hat{\boldsymbol{\mu}}_\tau^\zeta)^T(\hat{\mathbf{\Sigma}}_\tau^c)^{-1}(\hat{\boldsymbol{\mu}}_\tau^c - \hat{\boldsymbol{\mu}}_\tau^\zeta)\Bigg) \end{aligned} \tag{12}$$

in order to reduce the conflict between the transformed and the demonstrated joint distributions, where $|\cdot|$ and $\mathrm{Tr}(\cdot)$ represent the determinant and trace of matrix, respectively.

Note that (11) depends on $\boldsymbol{\theta}$, thus we can search for the optimal $\boldsymbol{\theta}$ that minimizes (11). To do so, many algorithms can be employed, e.g., gradient-based [16] and reward-weighted approaches [17]. We take the variant of policy improvement with path integrals [18] as an example. Assuming that we have defined an exploration distribution $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\theta}^{(0)}, \sigma^2\mathbf{I})$ with $\boldsymbol{\theta}^{(0)}$ being the initial parameters, we can sample $\boldsymbol{\theta}_l$ from this Gaussian distribution and measure the corresponding cost $c_l = J_{kl}(\boldsymbol{\theta}_l)$ according to (11). With L roll-outs, we can update $\boldsymbol{\theta}$ by $\boldsymbol{\theta}^{(1)} = \frac{\sum_{l=1}^{L} exp(-\alpha c_l)\boldsymbol{\theta}_l}{\sum_{l=1}^{L} exp(-\alpha c_l)}$, where $\alpha > 0$ is a constant. Similarly, by iteratively updating $\boldsymbol{\theta}$ until it converges, the optimal $\boldsymbol{\theta}^*$ can be determined. Finally, with the optimal $\boldsymbol{\theta}^*$ we can retrieve the transformed joint distribution by using (9). In this case, the original

objective (6) becomes

$$\tilde{J}(\mathbf{U}) = \sum_{\tau=t}^{t+h} \Big((\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^c)^T (\hat{\boldsymbol{\Sigma}}_\tau^c)^{-1} (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^c) + (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^\zeta)^T (\hat{\boldsymbol{\Sigma}}_\tau^\zeta)^{-1} (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^\zeta) \Big) + \sum_{\tau=t}^{t+h-1} \mathbf{u}_\tau^T \mathbf{R}_\tau \mathbf{u}_\tau. \tag{13}$$

Now, the objective (13) is represented in joint space. Similarly to the finite horizon LQR [9], [10], we can derive an analytic solution to (13).

D. Minimal Intervention Control with Competing Constraints

Let us write the system dynamics in joint space as

$$\boldsymbol{\zeta}_{t+1} = \underbrace{\begin{bmatrix} \mathbf{I} & \delta_t \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}}_{\mathbf{A}} \boldsymbol{\zeta}_t + \underbrace{\begin{bmatrix} \mathbf{0} \\ \delta_t \mathbf{I} \end{bmatrix}}_{\mathbf{B}} \mathbf{u}_t. \tag{14}$$

Then, given the current joint state $\boldsymbol{\zeta}_t$, we can predict the joint state sequence $\{\boldsymbol{\zeta}_\tau\}_{\tau=t}^{t+h}$ by iteratively using (14) (see [10] for details), i.e.,

$$\underbrace{\begin{bmatrix} \boldsymbol{\zeta}_t \\ \boldsymbol{\zeta}_{t+1} \\ \boldsymbol{\zeta}_{t+2} \\ \vdots \\ \boldsymbol{\zeta}_{t+h} \end{bmatrix}}_{\bar{\boldsymbol{\zeta}}} = \underbrace{\begin{bmatrix} \mathbf{I} \\ \mathbf{A} \\ \mathbf{A}^2 \\ \vdots \\ \mathbf{A}^h \end{bmatrix}}_{\bar{\mathbf{A}}} \boldsymbol{\zeta}_t + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{AB} & \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{A}^{h-1}\mathbf{B} & \mathbf{A}^{h-2}\mathbf{B} & \cdots & \mathbf{B} \end{bmatrix}}_{\bar{\mathbf{B}}} \underbrace{\begin{bmatrix} \mathbf{u}_t \\ \mathbf{u}_{t+1} \\ \vdots \\ \mathbf{u}_{t+h-1} \end{bmatrix}}_{\mathbf{U}}. \tag{15}$$

Furthermore, we can rewrite (13) into a compact form

$$\tilde{J}(\mathbf{U}) = (\bar{\boldsymbol{\zeta}} - \bar{\boldsymbol{\mu}}^c)^T (\bar{\boldsymbol{\Sigma}}^c)^{-1} (\bar{\boldsymbol{\zeta}} - \bar{\boldsymbol{\mu}}^c) + (\bar{\boldsymbol{\zeta}} - \bar{\boldsymbol{\mu}}^\zeta)^T (\bar{\boldsymbol{\Sigma}}^\zeta)^{-1} (\bar{\boldsymbol{\zeta}} - \bar{\boldsymbol{\mu}}^\zeta) + \mathbf{U}^T \bar{\mathbf{R}} \mathbf{U}. \tag{16}$$

with

$$\begin{aligned} \bar{\boldsymbol{\mu}}^c &= [(\hat{\boldsymbol{\mu}}_t^c)^T \ (\hat{\boldsymbol{\mu}}_{t+1}^c)^T \ \dots \ (\hat{\boldsymbol{\mu}}_{t+h}^c)^T]^T \\ \bar{\boldsymbol{\Sigma}}^c &= blockdiag(\hat{\boldsymbol{\Sigma}}_t^c, \hat{\boldsymbol{\Sigma}}_{t+1}^c, \dots, \hat{\boldsymbol{\Sigma}}_{t+h}^c) \\ \bar{\boldsymbol{\mu}}^\zeta &= [(\hat{\boldsymbol{\mu}}_t^\zeta)^T \ (\hat{\boldsymbol{\mu}}_{t+1}^\zeta)^T \ \dots \ (\hat{\boldsymbol{\mu}}_{t+h}^\zeta)^T]^T \\ \bar{\boldsymbol{\Sigma}}^\zeta &= blockdiag(\hat{\boldsymbol{\Sigma}}_t^\zeta, \hat{\boldsymbol{\Sigma}}_{t+1}^\zeta, \dots, \hat{\boldsymbol{\Sigma}}_{t+h}^\zeta) \\ \bar{\mathbf{R}} &= blockdiag(\mathbf{R}_t, \mathbf{R}_{t+1}, \dots, \mathbf{R}_{t+h-1}) \end{aligned}. \tag{17}$$

By substituting (15) (i.e., $\bar{\boldsymbol{\zeta}} = \bar{\mathbf{A}}\boldsymbol{\zeta}_t + \bar{\mathbf{B}}\mathbf{U}$) into (16) and calculating its derivative with respect to $\mathbf{U}$, we can derive the optimal $\mathbf{U}^*$ as

$$\mathbf{U}^* = \Big(\bar{\mathbf{B}}^T (\bar{\boldsymbol{\Sigma}}^c)^{-1} \bar{\mathbf{B}} + \bar{\mathbf{B}}^T (\bar{\boldsymbol{\Sigma}}^\zeta)^{-1} \bar{\mathbf{B}} + \bar{\mathbf{R}} \Big)^{-1} \Big(\bar{\mathbf{B}}^T (\bar{\boldsymbol{\Sigma}}^c)^{-1} (\bar{\boldsymbol{\mu}}^c - \bar{\mathbf{A}}\boldsymbol{\zeta}_t) + \bar{\mathbf{B}}^T (\bar{\boldsymbol{\Sigma}}^\zeta)^{-1} (\bar{\boldsymbol{\mu}}^\zeta - \bar{\mathbf{A}}\boldsymbol{\zeta}_t) \Big). \tag{18}$$

Note that if we only consider the joint space constraint in (13), i.e., $(\hat{\boldsymbol{\Sigma}}_\tau^c)^{-1} = \mathbf{0}$, the optimal solution (18) is exactly the solution to the finite horizon LQR [10]. The approach is summarized in Algorithm 1.

Algorithm 1 Minimal intervention control with competing constraints

1: Collect demonstrations $\{\{t_{n,m}, \boldsymbol{\xi}_{n,m}, \boldsymbol{\zeta}_{n,m}\}_{n=1}^N\}_{m=1}^M$
2: Extract probabilistic reference trajectories $\{\hat{\boldsymbol{\xi}}_t, \hat{\boldsymbol{\zeta}}_t\}_{t=1}^N$
3: Transform Cartesian constraint into joint space using (9)
4: Optimize null space parameters $\boldsymbol{\theta}$ so as to minimize (11)
5: Set h and $\{\mathbf{R}_t\}_{t=1}^{N-1}$
6: **for** $t = 1$ to $N-1$ **do**
7: Determine $\bar{\boldsymbol{\mu}}^c$, $\bar{\boldsymbol{\Sigma}}^c$, $\bar{\boldsymbol{\mu}}^\zeta$, $\bar{\boldsymbol{\Sigma}}^\zeta$ and $\bar{\mathbf{R}}$ using (17)
8: Compute the optimal command $\mathbf{U}^*$ by using (18)
9: Control the robot joint using $\mathbf{u}_t^*$
10: **end for**

IV. Dual Interpretation of Minimal Intervention Control

The optimal solution (18) can be interpreted from a dual perspective. Let us denote

$$\begin{aligned} \bar{\boldsymbol{\Sigma}}^d &= \Big((\bar{\boldsymbol{\Sigma}}^c)^{-1} + (\bar{\boldsymbol{\Sigma}}^\zeta)^{-1} \Big)^{-1} \quad \text{and} \\ \bar{\boldsymbol{\mu}}^d &= \bar{\boldsymbol{\Sigma}}^d \Big((\bar{\boldsymbol{\Sigma}}^c)^{-1} \bar{\boldsymbol{\mu}}^c + (\bar{\boldsymbol{\Sigma}}^\zeta)^{-1} \bar{\boldsymbol{\mu}}^\zeta \Big) \end{aligned}, \tag{19}$$

then (18) can be rewritten as

$$\mathbf{U}^* = \Big(\bar{\mathbf{B}}^T (\bar{\boldsymbol{\Sigma}}^d)^{-1} \bar{\mathbf{B}} + \bar{\mathbf{R}} \Big)^{-1} \Big(\bar{\mathbf{B}}^T (\bar{\boldsymbol{\Sigma}}^d)^{-1} (\bar{\boldsymbol{\mu}}^d - \bar{\mathbf{A}}\boldsymbol{\zeta}_t) \Big), \tag{20}$$

which is the solution of the finite horizon LQR defined by

$$\begin{aligned} \tilde{J}_d(\mathbf{U}) &= (\bar{\boldsymbol{\zeta}} - \bar{\boldsymbol{\mu}}^d)^T (\bar{\boldsymbol{\Sigma}}^d)^{-1} (\bar{\boldsymbol{\zeta}} - \bar{\boldsymbol{\mu}}^d) + \mathbf{U}^T \bar{\mathbf{R}} \mathbf{U} \\ &= \sum_{\tau=t}^{t+h} (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^d)^T (\hat{\boldsymbol{\Sigma}}_\tau^d)^{-1} (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^d) + \sum_{\tau=t}^{t+h-1} \mathbf{u}_\tau^T \mathbf{R}_\tau \mathbf{u}_\tau, \end{aligned} \tag{21}$$

where

$$\begin{aligned} \bar{\boldsymbol{\mu}}^d &= [(\hat{\boldsymbol{\mu}}_t^d)^T \ (\hat{\boldsymbol{\mu}}_{t+1}^d)^T \ \dots \ (\hat{\boldsymbol{\mu}}_{t+h}^d)^T]^T \\ \bar{\boldsymbol{\Sigma}}^d &= blockdiag(\hat{\boldsymbol{\Sigma}}_t^d, \hat{\boldsymbol{\Sigma}}_{t+1}^d, \dots, \hat{\boldsymbol{\Sigma}}_{t+h}^d) \\ \hat{\boldsymbol{\Sigma}}_\tau^d &= \Big((\hat{\boldsymbol{\Sigma}}_\tau^c)^{-1} + (\hat{\boldsymbol{\Sigma}}_\tau^\zeta)^{-1} \Big)^{-1} \\ \hat{\boldsymbol{\mu}}_\tau^d &= \hat{\boldsymbol{\Sigma}}_\tau^d \Big((\hat{\boldsymbol{\Sigma}}_\tau^c)^{-1} \hat{\boldsymbol{\mu}}_\tau^c + (\hat{\boldsymbol{\Sigma}}_\tau^\zeta)^{-1} \hat{\boldsymbol{\mu}}_\tau^\zeta \Big) \end{aligned}. \tag{22}$$

Thus, for the problem defined in (13), we can first calculate the product of the transformed joint distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^c, \hat{\boldsymbol{\Sigma}}_\tau^c)$ and the demonstrated joint distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_\tau^\zeta, \hat{\boldsymbol{\Sigma}}_\tau^\zeta)$, and subsequently employ the typical finite horizon LQR to calculate the optimal control command (joint acceleration in our case), leading to (20).

Note that similar insights have been pointed out in [4], [6], showing that Gaussian product can be used to mix competing constraints. Similarly, Gaussian product is also exploited in [10] when multiple constraints in different coordinate systems arise. However, these results are established at the pure trajectory level. Namely, they focus on generating an optimal trajectory that can address various trajectory constraints. In contrast, we here focus on a control problem and prove that (13) can be simplified by replacing different trajectory constraints by their Gaussian product (as shown in (21)), which offers an interesting insight when the previous work [4], [6] are combined with the finite LQR.

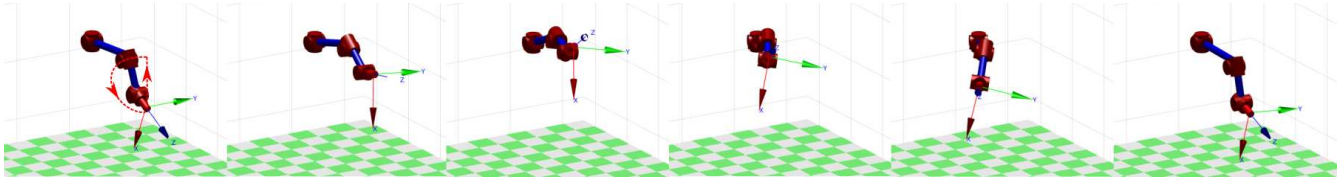

Fig. 2. Illustration of the writing task in a simulated robot, where the red curve depicts the robot motion direction.

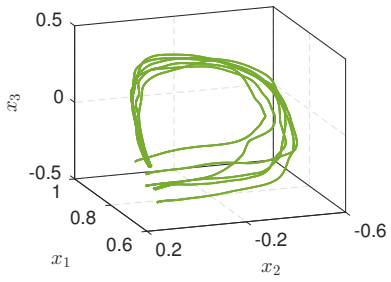

Fig. 3. Demonstrations of letter 'D' in the writing task.

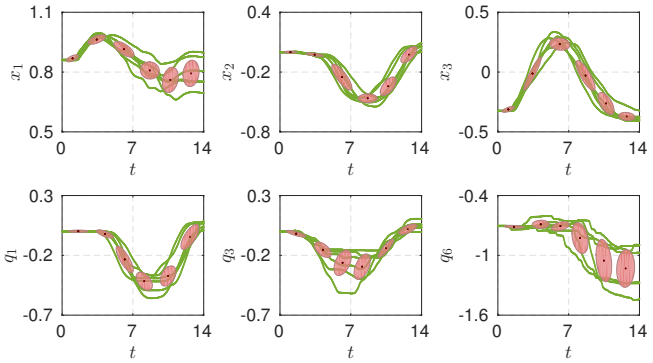

Fig. 4. GMM modeling of demonstrated trajectories, where the robot Cartesian trajectories as well as the first, third and sixth joint trajectories are shown. The ellipses represent Gaussian components.

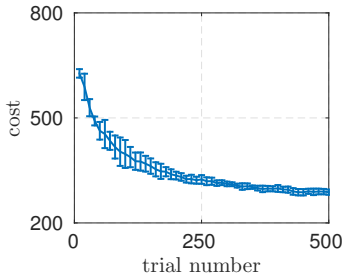

Fig. 5. The error-bar curve of KL-divergence between the transformed and original joint distributions when optimizing the null space parameters $\boldsymbol{\theta}$, where vertical bars denote the standard deviations.

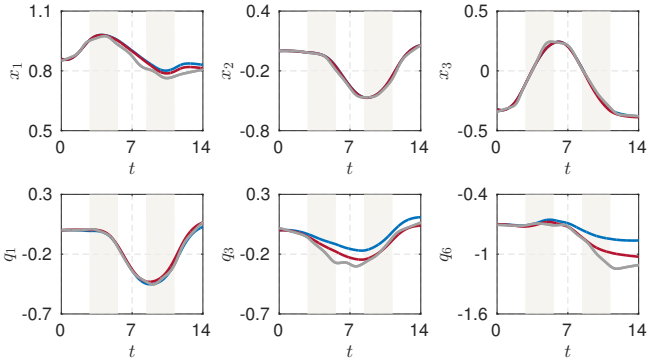

Fig. 6. The tracking Cartesian and joint trajectories with different methods under perturbations, where red and blue curves correspond to our approach and previous method, respectively. Gray curves represent the reference trajectories extracted from demonstrations.

V. Evaluations

In this section, we consider a writing task in a simulated 7-DoF Barrett robot, as depicted in Fig. 2. In order to verify the effectiveness of the null space exploration, we first compare our approach with previous work [4], [6] that ignores the robot null space (Section V-A). Then, we evaluate the tracking performance of both methods with/without external disturbances (Section V-B).

A. Evaluation of the Null-space Exploration

We collected six demonstrations for the writing task, as shown in Fig. 3 and Fig. 4, where the letter 'D' was written. It can be observed from Fig. 4 that the robot Cartesian trajectory in x_1 direction has a large variation in the time interval $8-14s$. Besides, the joint trajectories q_3 at $5-8s$ and q_6 at $8-14s$ are also less consistent. In the context of our framework, the trajectory segments with high variability in Cartesian or joint space correspond to small values of $\mathbf{Q}_\tau^\xi$ or $\mathbf{Q}_\tau^\zeta$ due to the large covariance, which hence allows for large tracking errors.

Following Algorithm 1, we use GMM to model the distributions of demonstrated trajectories in Cartesian and joint spaces, respectively. Subsequently, we employ GMR to extract probabilistic reference trajectory in each space. As can be seen in Fig. 4, the modeling result from GMM coincides with our observation of demonstration data, i.e., trajectory segments with small/large variations are associated with small/large covariances.

In order to show the effectiveness of the null space exploration, we transform the Cartesian constraints into joint space using our approach (where the normalized Gaussian functions are used as basis functions $\boldsymbol{\Phi}(t)$) and the previous method [4], [6], respectively. Specifically, we show the process of learning $\boldsymbol{\theta}$ in terms of the cost (11) in Fig. 5, where the converged cost is around 300. In contrast, the cost of previous methods [4], [6] is 614.55. Thus, the optimization of the robot null space effectively reduces the conflict between Cartesian and joint constraints.

TABLE I
EVALUATION OF DIFFERENT METHODS

	Without perturbation		With perturbation	
	Our approach	Previous method [4], [6]	Our approach	Previous method [4], [6]
c_e	3.18×10^3	1.44×10^4	5.83×10^3	2.07×10^4
c_u	9.99	9.53	84.31	83.93

B. Evaluations of Tracking Competing Trajectories

Now, we apply our method and the previous method [4], [6] to the tracking problem (6), where we consider the cases with and without external perturbations. The perturbations are added to the acceleration of seven joints directly during the time interval $2.8 - 5.6s$ and $8.4 - 11.2s$ with a fixed magnitude $0.2\ rad/s^2$. In order to compare both methods properly, we use the following two cost functions

$$\begin{aligned} c_e &= \sum_{\tau=1}^{N} (\boldsymbol{\xi}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\xi})^T (\hat{\boldsymbol{\Sigma}}_\tau^{\xi})^{-1} (\boldsymbol{\xi}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\xi}) + (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\zeta})^T (\hat{\boldsymbol{\Sigma}}_\tau^{\zeta})^{-1} (\boldsymbol{\zeta}_\tau - \hat{\boldsymbol{\mu}}_\tau^{\zeta}) \\ c_u &= \sum_{\tau=1}^{N} \mathbf{u}_\tau^T \mathbf{u}_\tau \end{aligned}, \tag{23}$$

where c_e and c_u correspond to the tracking cost and control effort. The comparison result is summarized in Table I, showing that with similar control efforts our approach achieves significantly smaller tracking errors. The simulated robot trajectories under perturbations are provided in Fig. 6. Both approaches are capable of generating trajectories that are near to the ones extracted from demonstrations. However, our method has smaller tracking errors in x_1, q_3 and q_6. Thus, we can conclude that by exploiting the robot null space towards reducing the conflict between competing constraints, the robot performance in term of (23) is indeed improved.

VI. CONCLUSIONS AND FUTURE WORK

We have studied the minimal intervention control associated with competing constraints and proved that the conflict between competing constraints can be reduced by minimizing a KL-divergence based objective. Moreover, we showed that the solution of the formulated control problem can be viewed as a classical finite horizon LQR. As shown in evaluations, by exploiting the robot null space we can achieve smaller tracking errors with similar control efforts.

Note that this paper only considers competing constraints that are extracted from demonstrations, which might prevent its application to the cases where significantly different trajectories from demonstrated examples are required. One extension could be the combination of trajectory adaptation approaches (e.g., various movement primitives [19], [20], [21] and hybrid imitation learning [7]) and the presented work. Also, we set the weight matrix $\mathbf{R}$ in (6) empirically, which is undesired for complicated systems. Thus, a further extension could be the learning of $\mathbf{R}$ (e.g., the energy-inspired estimation of $\mathbf{R}$ [22]). In addition, as done in previous works (e.g., [23]), the additional constraint can be viewed as a secondary objective in the robot null space, offering an alternative solution for combining Cartesian and joint trajectory constraints, and hence further investigation is still needed.

REFERENCES

[1] P. Pastor, H. Hoffmann, T. Asfour and S. Schaal, "Learning and generalization of motor skills by learning from demonstration," in *Proc. IEEE International Conference on Robotics and Automation*, 2009, pp. 763-768.

[2] Y. Huang, D. Büchler, O. Koc, B. Schölkopf and J. Peters, "Jointly learning trajectory generation and hitting point prediction in robot table tennis," in *Proc. IEEE International Conference on Humanoid Robots*, 2016, pp. 650-655.

[3] Y. Zhou and T. Asfour, "Task-oriented generalization of dynamic movement primitive," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2017, 3202 - 3209.

[4] J. Silvério, S. Calinon, L. Rozo and D. G. Caldwell, "Learning task priorities from demonstrations," *arXiv*:1707.06791v2, 2017.

[5] Y. Huang, B. Schölkopf and J. Peters, "Learning optimal striking points for a ping-pong playing robot," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2015, pp. 4587-4592.

[6] S. Calinon and A. Billard, "Statistical learning by imitation of competing constraints in joint space and task space," *Advanced Robotics*, vol. 23, pp. 2059-2076, 2009.

[7] Y. Huang, J. Silvério, L. Rozo, and D. G. Caldwell, "Hybrid probabilistic trajectory optimization using null-space exploration," in *Proc. IEEE International Conference on Robotics and Automation*, 2018.

[8] J. R. Medina, D. Lee and S. Hirche, "Risk-sensitive optimal feedback control for haptic assistance," in *Proc. IEEE International Conference on Robotics and Automation*, 2012, pp. 1025-1031.

[9] S. Calinon, D. Bruno and D. G. Caldwell, "A task-parameterized probabilistic model with minimal intervention control," in *Proc. IEEE International Conference on Robotics and Automation*, 2014, pp. 3339-3344.

[10] S. Calinon, "A tutorial on task-parameterized movement learning and retrieval," *Intelligent Service Robotics*, vol. 9, no. 1, pp. 1-29, 2016.

[11] M. Muhlig, M. Gienger, S. Hellbach, J. J. Steil and C. Goerick, "Task-level imitation learning using variance-based movement optimization," in *Proc. IEEE International Conference on Robotics and Automation*, 2009, pp. 1177-1184.

[12] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. Appendix A.2, Appendix A.5, Cambridge: MIT press, 2006.

[13] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79-86, 1951.

[14] J. Peters, K. Mülling and Y. Altun, "Relative entropy policy search," in *Proc. AAAI*, 2010, pp. 1607-1612.

[15] S. Levine and P. Abbeel, "Learning neural network policies with guided policy search under unknown dynamics," in *Proc. Advances in Neural Information Processing Systems*, 2014, pp. 1071-1079.

[16] J. Peters and S. Schaal, "Policy gradient methods for robotics," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 2219-2225.

[17] E. Theodorou, J. Buchli and S. Schaal, "A generalized path integral control approach to reinforcement learning," *Journal of Machine Learning Research*, vol. 11, pp. 3137-3181, 2010.

[18] F. Stulp and O. Sigaud, "Robot skill learning: from reinforcement learning to evolution strategies," *Journal of Behavioral Robotics*, vol. 4, no. 1, pp. 49-61, 2013.

[19] Y. Huang, L. Rozo, J. Silvério and D. G. Caldwell. "Kernelized movement primitives," *arXiv*:1708.08638v2, 2017.

[20] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor and S. Schaal, "Dynamical movement primitives: learning attractor models for motor behaviors," *Neural Computation*, vol. 25, no. 2, pp. 328-373, 2013.

[21] A. Paraschos, C. Daniel, J. Peters, and G. Neumann, "Probabilistic movement primitives," in *Proc. Advances in Neural Information Processing Systems*, 2616-2624, 2013.

[22] J. Silvério, Y. Huang, L. Rozo, and D. G. Caldwell, "An uncertainty-aware minimal intervention control strategy learned from demonstrations," in *Proc. IEEE International Conference on Intelligent Robots and Systems*, 2018.

[23] J. Hollerbach and K. Suh, "Redundancy resolution of manipulators through torque optimization," in *Proc. IEEE International Conference on Robotics and Automation*, 1985, pp. 1016-1021.