

Learning to Avoid Obstacles With Minimal Intervention Control

Original

Learning to Avoid Obstacles With Minimal Intervention Control / Duan, A; Camoriano, R; Ferigo, D; Huang, Y; Calandriello, D; Rosasco, L; Pucci, D. - In: FRONTIERS IN ROBOTICS AND AI. - ISSN 2296-9144. - ELETTRONICO. - 7:(2020). [10.3389/frobt.2020.00060]

Availability:

This version is available at: 11583/2982137 since: 2023-09-13T20:04:27Z

Publisher:

Frontiers

Published

DOI:10.3389/frobt.2020.00060

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)



Learning to Avoid Obstacles With Minimal Intervention Control

Anqing Duan^{1,2*}, Raffaello Camoriano³, Diego Ferigo^{1,4}, Yanlong Huang⁵,
Daniele Calandriello³, Lorenzo Rosasco^{2,3,6} and Daniele Pucci¹

¹ Dynamic Interaction Control Lab, Italian Institute of Technology, Genoa, Italy, ² DIBRIS, Università degli Studi di Genova, Genoa, Italy, ³ Laboratory for Computational and Statistical Learning (IIT@MIT), Istituto Italiano di Tecnologia and Massachusetts Institute of Technology, Cambridge, MA, United States, ⁴ Machine Learning and Optimization Group, School of Computer Science, University of Manchester, Manchester, United Kingdom, ⁵ Department of Advanced Robotics, Istituto Italiano di Tecnologia, Genoa, Italy, ⁶ Machine Learning Genoa (MaLGa) Center, Università di Genova, Genoa, Italy

OPEN ACCESS

Edited by:

Bojan Nemeč,
Institut Jožef Stefan (IJS), Slovenia

Reviewed by:

Charalampos P. Bechlioulis,
National Technical University of
Athens, Greece
Leon Žitjač,
Institut Jožef Stefan (IJS), Slovenia

*Correspondence:

Anqing Duan
anqing.duan@iit.it

Specialty section:

This article was submitted to
Robotic Control Systems,
a section of the journal
Frontiers in Robotics and AI

Received: 13 May 2019

Accepted: 08 April 2020

Published: 28 May 2020

Citation:

Duan A, Camoriano R, Ferigo D,
Huang Y, Calandriello D, Rosasco L
and Pucci D (2020) Learning to Avoid
Obstacles With Minimal Intervention
Control. *Front. Robot. AI* 7:60.
doi: 10.3389/frobt.2020.00060

Programming by demonstration has received much attention as it offers a general framework which allows robots to efficiently acquire novel motor skills from a human teacher. While traditional imitation learning that only focuses on either Cartesian or joint space might become inappropriate in situations where both spaces are equally important (e.g., writing or striking task), hybrid imitation learning of skills in both Cartesian and joint spaces simultaneously has been studied recently. However, an important issue which often arises in dynamical or unstructured environments is overlooked, namely how can a robot avoid obstacles? In this paper, we aim to address the problem of avoiding obstacles in the context of hybrid imitation learning. Specifically, we propose to tackle three subproblems: (1) designing a proper potential field so as to bypass obstacles, (2) guaranteeing joint limits are respected when adjusting trajectories in the process of avoiding obstacles, and (3) determining proper control commands for robots such that potential human-robot interaction is safe. By solving the aforementioned subproblems, the robot is capable of generalizing observed skills to new situations featuring obstacles in a feasible and safe manner. The effectiveness of the proposed method is validated through a toy example as well as a real transportation experiment on the iCub humanoid robot.

Keywords: programming by demonstration, reinforcement learning, obstacle avoidance, humanoid robots, minimal intervention control

1. INTRODUCTION

Over the past few years, there has been growing demand for bringing robots from industrial manufacturing lines to human-centered scenarios thanks to ever evolving sensors, actuators, and processors. Increasing computational power also gives rise to novel control and learning algorithms. Nevertheless, contrary to the high maturity of industrial robots and relatively simple service robots, how to deploy general purpose robots, such as humanoid robots into cluttered environments still remains a formidable challenge. Conceivably, before humanoids can successfully operate in daily-life settings, a series of challenges need to be confronted. One of the major concerns for robots to operate outside laboratory environments is obstacle avoidance. Indeed, obstacle avoidance represents a necessary capability for robots to become more autonomous, flexible and

safer in order to cope with complex working environments. In addition, as humanoid robots could sometimes be expected to work alongside human beings, classical high-gain feedback control shall be deprecated in case human-robot interaction becomes dangerous. Hence, compliance is a highly desired skill in this case, as the incorporation of variable impedance skills into the robot controller allows for a safe physical human-robot interaction (Abu-Dakka et al., 2018).

In view of high-dimensional state and action spaces of humanoid robots, a user-friendly method to endow them with various skills is Programming by Demonstration (PbD), also known as imitation learning (Billard et al., 2008). The complete procedure of PbD typically consists of a demonstration phase, where robots are shown the desired behavior, and a reproduction phase, where robots are required to reproduce the learned skills, typically with the help of movement primitives (Ijspeert et al., 2013; Huang et al., 2019). Under the framework of PbD, humanoid robots are able to efficiently acquire novel motor skills from demonstrations of a human teacher. Following this paradigm, many successful results have been achieved, such as peg-in-hole task (Abu-Dakka et al., 2014), cleaning task (Duan et al., 2019), table tennis (Huang et al., 2016), etc.

Traditional imitation learning that only focuses on either Cartesian or joint space might become inappropriate in situations where both spaces are equally important, such as in writing or striking tasks. Consequently, hybrid imitation learning of skills in both Cartesian and joint spaces simultaneously has recently emerged. In order to further generalize the applicability of hybrid imitation learning, this paper considers the integration of obstacle avoidance in the context of hybrid imitation learning such that robots can reproduce the learned skills in a broader range of situations. Specifically, we consider the following three aspects: (1) designing a proper potential field so as to bypass obstacles. As a common technique for obstacle avoidance within PbD, the potential field formulation as well as its hyperparameters play an important role in realizing obstacle avoidance. In order to compare the performance of different potential fields, a novel imitation metric is proposed. Moreover, a kernel-based reinforcement learning algorithm is employed to determine the hyperparameters of the chosen potential field; (2) guaranteeing that joint limits are respected when adjusting trajectories in the process of avoiding obstacles. During obstacle avoidance, joint trajectories are usually modified according to the effects of the potential field. Therefore, the evolution of joint trajectories shall be constrained by bounding them within the allowable range; (3) determining proper control commands for robots such that potential human-robot interaction is safe. To do so, We propose to control the robot with a minimal intervention controller rooted in linear quadratic tracking.

The rest of the paper is organized as follows: section 2 reviews the previous work related to our problem. Section 3 presents the proposed framework for learning to avoid obstacles with minimal intervention control. Subsequently, section 4 reports the results of the toy example as well as the experiments on the iCub humanoid robot to show the effectiveness of the proposed method. Finally, conclusions and future routes of research are given in section 5.

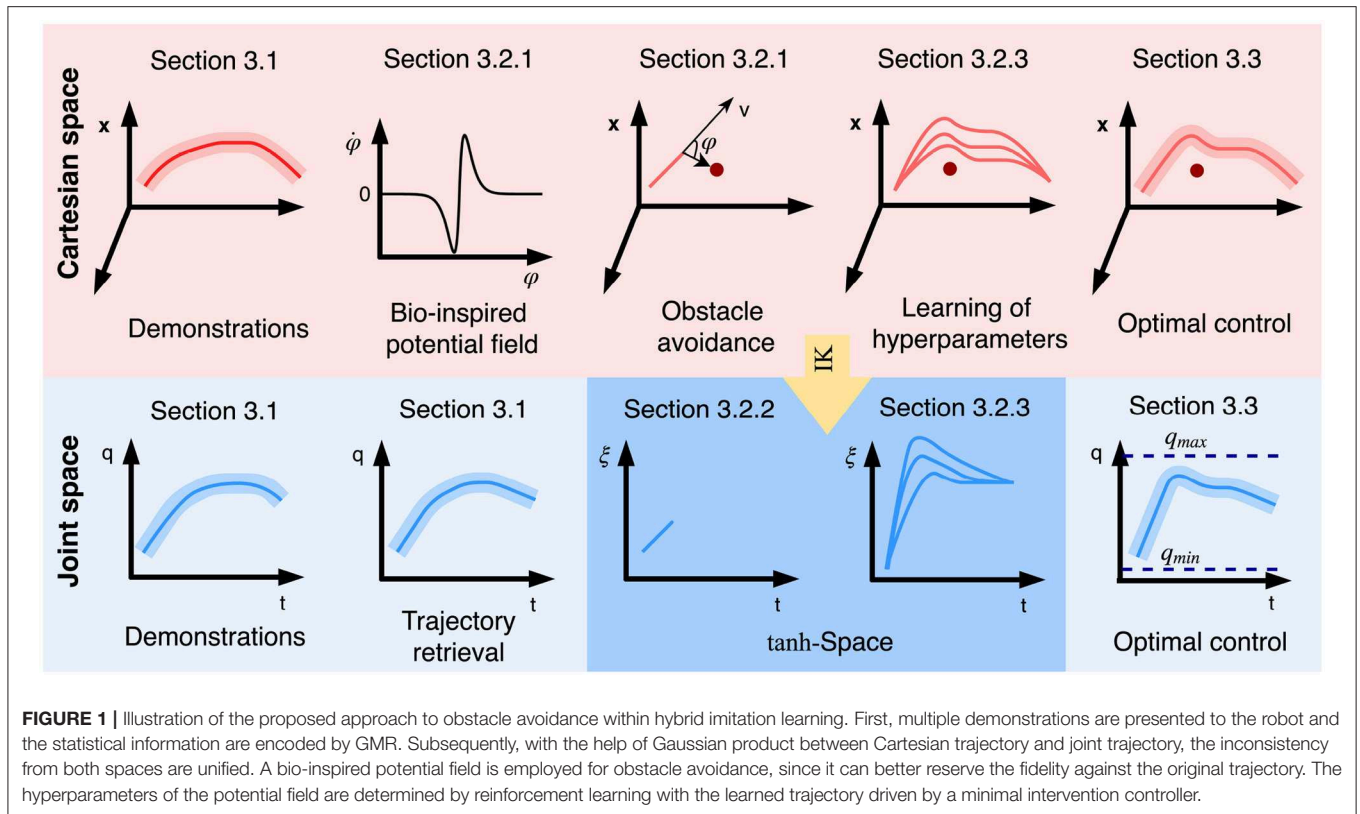
2. RELATED WORK

In general, obstacle avoidance is a classical topic. Due to its great significance, it has been extensively studied in a broad range of fields not only limited to robotics but also computer graphics, computer aided design, urban planning, etc. Among the considerable works dedicated to the topic, obstacle avoidance can be roughly classified into two categories: motion planning and reactive methods.

Sampling-based motion planning algorithms normally rely on planners, such as Probabilistic Roadmap (PRM) (Kavraki et al., 1996) and Rapidly-exploring Random Tree (RRT) algorithms (LaValle, 1998), along with their numerous extensions. In order to facilitate collision detection of the samples, polyhedrons are usually used as proxies for robots and obstacles. Collision avoidance strategies elicited from sampling-based motion planning usually could generate globally optimal trajectories, but become computationally expensive and time consuming in the case of high-dimensional multi-body problems or narrow-passage problems (Ruan et al., 2018). Optimization-based techniques can also be employed for obstacle avoidance. The collision-free trajectory can be obtained by optimizing the cost function formulated by a combination of obstacle cost and other indexes, such as smoothness. Various strategies for optimization could be applied to motion planning as well. For example, Zucker et al. (2013) presented a trajectory optimization procedure based on covariant descent. However, considering that gradient-based methods could get stuck in local optima, Kalakrishnan et al. (2011) proposed to instead use a derivative-free stochastic optimization approach to motion planning. The optimized trajectory is obtained by rolling out a series of noisy trajectories and the candidate solution is updated with the received cost with no gradient information required in the process. Also, in order to speed up the optimization process, Schulman et al. (2014) proposed to find collision-free trajectories using sequential convex optimization where collision is penalized with a hinge loss.

By contrast, reactive methods can make sure that robots can behave in response to the sensed obstacles in real time. Yet, the limitations lie in the design of priority assignment between the ongoing tasks as well as the obstacle avoidance task. Therefore, the solution is usually satisfying local conditions and thus suboptimal. In addition, there are also stability issues regarding reactive methods, as identified by Koren and Borenstein (1991).

In the framework of PbD, obstacle avoidance is usually realized with the help of potential fields, i.e., collision-free movement is generated by a repellent force obtained from a gradient of a potential field centered around the obstacle (Khatib, 1986). Kim and Khosla (1992) proposed a new formulation of the artificial potential field for obstacle avoidance using harmonic functions. A clear advantage of harmonic functions is that it can completely eliminate local minima in a cluttered environment. In the spirit of harmonic potential functions, within the context of dynamical-system-based robot control, Khansari-Zadeh and Billard (2012) proposed a real-time obstacle avoidance approach that can steer robot motions generated by the dynamical system. Since such modification happens locally, the dynamical system



(DS) can still be kept globally stable, i.e., all trajectories can reach the target point. In addition, the method can handle multiple obstacles without changing the equilibrium of the original dynamics. Huber et al. (2019) further extended such method to the case of moving convex and star-shaped concave obstacles. The impenetrability of the obstacles' hull and asymptotic stability at a final goal location was proven by contraction theory.

Compared with previous work on obstacle avoidance within PbD, the contribution of our proposed method focuses on tackling the aforementioned three subproblems. To address problem (1), a novel imitation metric is provided such that the performance of different potential fields can be quantified. Based on such imitation metric, we can evaluate similarity between the trajectories modified due to obstacle avoidance as well as the original demonstrated trajectories. The hyperparameters of the chosen potential field are determined by a reinforcement learning algorithm. To solve problem (2), we parameterize the joint space into exogenous states using the hyperbolic tangent function. We show that the proposed method can guarantee that the evolution of joint trajectories is always bounded within the specified range. As for problem (3), we employ a minimal intervention control strategy based on linear quadratic tracking. The illustration of the proposed method is shown in **Figure 1**. In addition, a flowchart to summarize the whole procedure is shown in **Figure 2**.

3. PROPOSED APPROACH

In this section we propose an obstacle avoidance approach within PbD that aims to preserve the demonstrated trajectories. Our

obstacle avoidance strategy is devised based on the principle of artificial potential field. With the goal of preserving the demonstrated trajectories, we propose to use a bio-inspired potential field called Fajen potential field proposed by Fajen and Warren (2003). The Fajen potential field is built upon the empirical evidence of how humans steer their motion for obstacle avoidance. The hyperparameters of the chosen potential field are determined by a reinforcement learning algorithm. Although the Fajen field has been used in previous works (Hoffmann et al., 2009; Rai et al., 2014), its merit over others was not elaborated explicitly. To this end, we contribute quantitative evidence to benchmark different types of potential fields and show that the Fajen potential field indeed outperforms others. Furthermore, we propose to rely on a novel imitation metric rather than RMSE to evaluate the imitation fidelity. The suggested imitation metric is based on the technique of curve similarity analysis (Dryden, 2014). Moreover, given that trajectories will be modified by the potential field unpredictably during obstacle avoidance, there are possibilities that joint trajectory evolution could exceed the allowable range. To address this issue, we employ Constrained Dynamic Movement Primitives (CDMPs), recently developed by Duan et al. (2018) so as to ensure joint trajectories are always bounded within the specified range.

3.1. Trajectory Retrieval From Multiple Demonstrations

In order to endow robots with variable impedance skills, a minimal intervention controller is usually used for tracking the reference trajectories. Multiple demonstrations to the robot

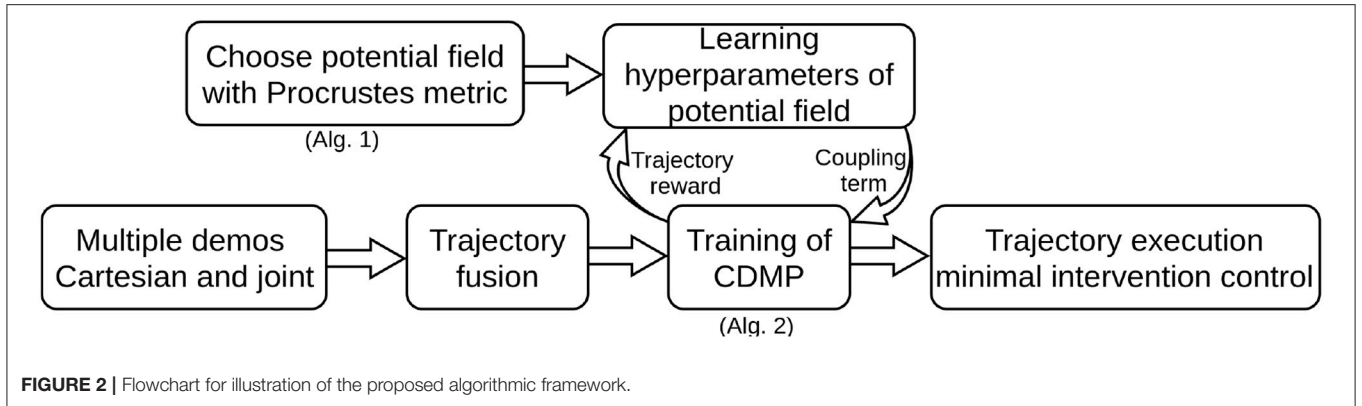


FIGURE 2 | Flowchart for illustration of the proposed algorithmic framework.

are required to capture the statistical information underlying the reference trajectories. Assume that both Cartesian trajectory $\tau_t^x = [\mathbf{x}_t^\top \dot{\mathbf{x}}_t^\top]^\top$ and joint trajectory $\tau_t^j = [\mathbf{q}_t^\top \dot{\mathbf{q}}_t^\top]^\top$ as well as their corresponding time index t are recorded. For M demonstrations, each one having length T , we denote the collected dataset as $\{\{t_{n,m}, \tau_{t,m}^x, \tau_{t,m}^j\}_{t=1}^T\}_{m=1}^M$. In order to extract the probabilistic properties from the multiple teachings, several techniques can be employed, such as Gaussian Mixture Models (GMM) or Hidden Markov Models (HMM) (Calinon and Lee, 2016). As an example, here we employ GMM with H components to encode the raw training data as GMM is one of the most mature probabilistic approaches for modeling multiple demonstrations. Without loss of generality, we use τ to denote either τ^x or τ^j . GMM first estimates the joint probability distribution with priors $\pi_h, h = 1, \dots, H$ satisfying $\sum_{h=1}^H \pi_h = 1$:

$$p(t, \tau) = \sum_{h=1}^H \pi_h \mathcal{N}(\mu_h, \Sigma_h), \quad (1)$$

where

$$\mu_h = \begin{bmatrix} \mu_{t,h} \\ \mu_{\tau,h} \end{bmatrix}, \quad (2)$$

$$\Sigma_h = \begin{bmatrix} \Sigma_{tt,h} & \Sigma_{t\tau,h} \\ \Sigma_{\tau t,h} & \Sigma_{\tau\tau,h} \end{bmatrix}. \quad (3)$$

Furthermore, Gaussian Mixture Regression (GMR) is employed to retrieve the probabilistic trajectory (Calinon and Lee, 2016). The corresponding output with respect to a query point t is formulated by a conditional probability distribution:

$$\hat{\tau}(t) \sim \sum_{h=1}^H w_h(t) \mathcal{N}(\hat{\mu}_h(t), \hat{\Sigma}_h(t)), \quad (4)$$

where $w_h(t)$ are the activation functions defined as

$$w_h(t) = \frac{\pi_h \mathcal{N}(t | \mu_{t,h}, \Sigma_{tt,h})}{\sum_{i=1}^H \pi_i \mathcal{N}(t | \mu_{t,i}, \Sigma_{tt,i})}, \quad (5)$$

with

$$\hat{\mu}_h(t) = \mu_{\tau,h} + \Sigma_{\tau t,h} \Sigma_{tt,h}^{-1} (t - \mu_{t,h}), \quad (6)$$

$$\hat{\Sigma}_h(t) = \Sigma_{\tau\tau,h} - \Sigma_{\tau t,h} \Sigma_{tt,h}^{-1} \Sigma_{t\tau,h}. \quad (7)$$

Note that (4) is usually approximated by a unimodal output distribution for robot control. By resorting to the law of total mean and variance, the approximated normal distribution $\mathcal{N}(\hat{\mu}_t^u, \hat{\Sigma}_t^u)$ can be derived as

$$\begin{aligned} \hat{\mu}_t^u &= \sum_{h=1}^H w_h(t) \hat{\mu}_h(t), \\ \hat{\Sigma}_t^u &= \sum_{h=1}^H w_h(t) (\hat{\Sigma}_h(t) + \hat{\mu}_h(t) \hat{\mu}_h(t)^\top) - \hat{\mu}_t^u \hat{\mu}_t^{u\top}. \end{aligned} \quad (8)$$

After applying GMR, we obtain $\hat{\tau}^x \sim \mathcal{N}(\hat{\mu}^{\tau x}, \hat{\Sigma}^{\tau x})$ for Cartesian trajectories where $\hat{\mu}^{\tau x} = [\hat{\mu}^{x\top} \hat{\mu}^{x\dot{\top}}]^\top$, and $\hat{\tau}^j \sim \mathcal{N}(\hat{\mu}^{\tau j}, \hat{\Sigma}^{\tau j})$ for joint trajectories where $\hat{\mu}^{\tau j} = [\hat{\mu}^{j\top} \hat{\mu}^{j\dot{\top}}]^\top$. One issue arising here is that there is inconsistency emerging between Cartesian and joint constraints due to multiple demonstrations. Such phenomena are usually referred to as *competing constraints* in the literature (Calinon and Billard, 2008). In order to unify the constraints from both spaces, we employ the Gaussian product for the fusion of Cartesian and joint trajectories as in Calinon and Billard (2008). To obtain the corresponding joint trajectory \mathbf{q} that satisfies the corresponding Cartesian constraints \mathbf{x} , a Jacobian-based inverse kinematics technique is employed:

$$\begin{aligned} \mathbf{q}_t^x &= \mathbf{q}_{t-1}^x + \mathbf{J}^\dagger (\mathbf{x}_t - \mathbf{x}_{t-1}) + \mathbf{J}^N (\mathbf{q}_t - \mathbf{q}_{t-1}), \\ \dot{\mathbf{q}}_t^x &= \mathbf{J}^\dagger \dot{\mathbf{x}}_t + \mathbf{J}^N \dot{\mathbf{q}}_t, \end{aligned} \quad (9)$$

where $\mathbf{J}^\dagger = \mathbf{J}^\top (\mathbf{J}\mathbf{J}^\top)^{-1}$ denotes the Moore-Penrose pseudo inverse of \mathbf{J} and $\mathbf{J}^N = \mathbf{I} - \mathbf{J}^\dagger \mathbf{J}$ is a null-space matrix that projects the additional secondary task into the null space of robot movement. In this case, the joint trajectories $\hat{\tau}^j \sim \mathcal{N}(\hat{\mu}^{\tau j}, \hat{\Sigma}^{\tau j})$ retrieved from demonstrations are set as the secondary task and thus represent null-space movement. Upon transformation from Cartesian to joint space, the obtained probabilistic joint trajectory

also satisfies Gaussian distribution $\mathbf{q}^C \sim \mathcal{N}(\boldsymbol{\mu}^C, \boldsymbol{\Sigma}^C)$ with mean $\boldsymbol{\mu}^C$ and variance $\boldsymbol{\Sigma}^C$ given by:

$$\begin{aligned}\boldsymbol{\mu}^C &= \tilde{\mathbf{J}}^\dagger \hat{\boldsymbol{\mu}}^{\tau x} + \tilde{\mathbf{J}}^N \hat{\boldsymbol{\mu}}^{\tau j} + \boldsymbol{\Lambda}_{t-1}, \\ \boldsymbol{\Sigma}^C &= \tilde{\mathbf{J}}^\dagger \hat{\boldsymbol{\Sigma}}^{\tau x} \tilde{\mathbf{J}}^{\dagger\top} + \tilde{\mathbf{J}}^N \hat{\boldsymbol{\Sigma}}^{\tau j} \tilde{\mathbf{J}}^{N\top},\end{aligned}\quad (10)$$

where $\boldsymbol{\Lambda}_{t-1} = [[\boldsymbol{\mu}_{t-1}^C - \mathbf{J}^\dagger \hat{\boldsymbol{\mu}}_{t-1}^x - \mathbf{J}^N \hat{\boldsymbol{\mu}}_{t-1}^j]^\top \mathbf{0}^\top]^\top$, $\tilde{\mathbf{J}}^\dagger = \text{diagblock}(\mathbf{J}^\dagger, \mathbf{J}^\dagger)$ and $\tilde{\mathbf{J}}^N = \text{diagblock}(\mathbf{J}^N, \mathbf{J}^N)$. To address the competing constraints, the final probabilistic trajectory $\hat{\mathbf{q}} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ is retrieved by fusion of Cartesian and joint constraints with the help of the Gaussian product (Calinon and Billard, 2008). That is, $\hat{\mathbf{q}} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) \propto \mathcal{N}(\boldsymbol{\mu}^C, \boldsymbol{\Sigma}^C) \times \mathcal{N}(\hat{\boldsymbol{\mu}}^{\tau j}, \hat{\boldsymbol{\Sigma}}^{\tau j})$ with mean $\hat{\boldsymbol{\mu}}$ and variance $\hat{\boldsymbol{\Sigma}}$ given by Rasmussen (2003):

$$\begin{aligned}\hat{\boldsymbol{\mu}} &= \bar{\boldsymbol{\Sigma}}((\boldsymbol{\Sigma}^C)^{-1} \boldsymbol{\mu}^C + (\hat{\boldsymbol{\Sigma}}^{\tau j})^{-1} \hat{\boldsymbol{\mu}}^{\tau j}), \\ \hat{\boldsymbol{\Sigma}} &= ((\boldsymbol{\Sigma}^C)^{-1} + (\hat{\boldsymbol{\Sigma}}^{\tau j})^{-1})^{-1}.\end{aligned}\quad (11)$$

By executing the retrieved probabilistic trajectory, robots are able to satisfy both Cartesian and joint constraints simultaneously.

3.2. Learning to Avoid Obstacles

Once the probabilistic trajectories are retrieved, the robot will be required to track the reference trajectories in presence of obstacles. We present a novel imitation metric other than the simple sum of squares in section 3.2.1 in order to compare the performance of different potential fields. The joint limit avoidance issue is addressed in section 3.2.2. Subsequently, section 3.2.3 formulates the search of the optimal potential field hyperparameters in terms of a Reinforcement Learning problem.

3.2.1. Imitation Metric for Potential Field

Since we expect robots to behave like humans during the process of obstacle avoidance, our choice for the potential field is the Fajen potential field, which is derived from a bio-inspired perspective (Fajen and Warren, 2003). The basic idea behind Fajen field is to first calculate the angle between the current velocity and the direction toward the obstacle. Given this angle, the method determines how much to change the steering direction in order to keep away from the obstacle. The steering effect from the Fajen potential field is used to design the coupling term:

$$\mathbf{p}(\mathbf{x}) = \gamma \mathbf{R} \dot{\mathbf{x}} \varphi e^{-\beta \varphi}, \quad (12)$$

with

$$\varphi = \arccos\left(\frac{(\mathbf{o} - \mathbf{x})^\top \dot{\mathbf{x}}}{\|\mathbf{o} - \mathbf{x}\| \|\dot{\mathbf{x}}\|}\right), \quad (13)$$

$$\mathbf{r} = (\mathbf{o} - \mathbf{x}) \times \dot{\mathbf{x}}, \quad (14)$$

where \mathbf{R} is a rotation matrix with axis \mathbf{r} and rotation angle $\pi/2$, \mathbf{o} denotes the position of the obstacle in Cartesian space and φ is the angle between the velocity of the end-effector $\dot{\mathbf{x}}$ and the vector $\mathbf{o} - \mathbf{x}$, which is always positive.

As there are a number of other potential fields that can also be used for obstacle avoidance within PbD, an interesting

issue arising is how to compare the performance of different potential fields. In order to evaluate the reproduction quality of the trajectory, imitation metrics are required. Traditionally, imitation metrics are mainly formulated as a weighted sum of squares of differences between the reproduced and the demonstrated trajectories, where the weights usually come from the variance matrix across multiple demonstrations (Billard et al., 2008; Calinon and Lee, 2016; Huang et al., 2018). Such evaluation metric is improper in the case of obstacle avoidance since the trajectory shapes are changed much more severely with respect to straightforward reproduction.

Here, we employ a novel perspective on the formalism of imitation metrics, such that the effects of different potential fields can be fairly evaluated. Specifically, we propose to formulate the imitation similarity metric from the perspective of curve similarity analysis (Dryden, 2014). In general, the technique of curve similarity analysis has a wide range of applications, such as signal alignment, DNA matching, signature comparison, etc. and is very pertinent to our situation (Mitchel et al., 2018). As there are a number of curve similarity analysis methods, the Procrustes distance is used for our case as an example (Dryden, 2014).

Overall, the Procrustes distance facilitates shape analysis by removing relative translational, scaling, and rotational components. Here we consider how to calculate the Procrustes distance $d(X^1, X^2)$ between two trajectories $X^1 = \{[x_k^1, y_k^1, z_k^1]^\top\}_{k=1}^K$ and $X^2 = \{[x_k^2, y_k^2, z_k^2]^\top\}_{k=1}^K$ with K the length. First, the translational component will be removed. To this end, all the trajectory points are translated as a whole such that the mean value of all the points coincide with the origin point. The mean value \bar{X}^i of all the points for trajectory i is calculated as

$$\bar{X}^i = \frac{\sum_k X_k^i}{K}. \quad (15)$$

Similarly, the scaling component is removed by normalizing the root mean square distance between the points and the origin point. The scaling factor s^i of trajectory i is calculated as

$$s^i = \sqrt{\frac{\sum_k (X_k^i - \bar{X}^i)^2}{K}}. \quad (16)$$

The scale of trajectory i is normalized by

$$\tilde{X}_k^i = \frac{X_k^i - \bar{X}^i}{s^i} \quad (17)$$

The removal of the rotation effect is not straightforward because the calculation of the optimal rotation matrix \mathbf{W} requires to solve an optimization problem

$$\mathbf{W} = \arg \max_{\mathbf{Y}} \sum_k \|\mathbf{Y} \tilde{X}_k^1 - \tilde{X}_k^2\|^2, \quad (18)$$

where $\|\cdot\|$ denotes the Euclidean norm of a vector. It can be verified that the optimal rotation matrix is given by Dryden (2014)

$$\mathbf{W} = \mathbf{U} \boldsymbol{\Sigma}' \mathbf{V}^\top, \quad (19)$$

where $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ is the Singular Value Decomposition (SVD) of the matrix $\sum_k \tilde{X}_k^2 \tilde{X}_k^{1\top}$. In order to make \mathbf{W} a valid rotation matrix, i.e., $\det(\mathbf{W}) > 0$ where $\det(\cdot)$ denotes the determinant of a matrix, $\mathbf{\Sigma}$ is modified into $\mathbf{\Sigma}'$ by replacing its smallest singular value with the sign of $\det(\mathbf{U}\mathbf{V}^\top)$, and the other singular values with 1. Finally, the Procrustes distance is given by

$$d(X^1, X^2) = \sqrt{\sum_k \|\tilde{X}_k^1 - \mathbf{W}\tilde{X}_k^2\|^2}. \quad (20)$$

The process to calculate the Procrustes between two trajectories is summarized in Algorithm 1.

Algorithm 1: Procrustes distance between trajectories

Input: Trajectories X^1 and X^2

Output: Procrustes distance $d(X^1, X^2)$

- 1: Compute \tilde{X}^i as per (15)
 - 2: Compute s^i as per (16)
 - 3: **for** $k = 1$ to K **do**
 - 4: $X_k^i \leftarrow (X_k^i - \tilde{X}^i)/s^i$
 - 5: **end for**
 - 6: $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top \leftarrow \text{SVD}(\sum_k X_k^2 X_k^{1\top})$
 - 7: **if** $\det(\mathbf{U}\mathbf{V}) < 0$ **then**
 - 8: $\min(\mathbf{\Sigma}) \leftarrow \text{sign}(\det(\mathbf{U}\mathbf{V}))$
 - 9: $\text{nonmin}(\mathbf{\Sigma}) \leftarrow 1$
 - 10: **end if**
 - 11: Compute d as per (20)
 - 12: **return** d
-

3.2.2. Joint Limit Avoidance

Once a suitable type of potential field has been chosen, we are ready to modify robot trajectories for obstacle avoidance. One problem here is how to guarantee that joint limits are respected. It can be conceived that the modified trajectories are very susceptible to the strength of the potential field and a strong field could drive joint trajectories out of the allowable range. To cope with such issue, we will drive the robot's trajectories using our recently developed Constrained Dynamic Movement Primitives (CDMPs) (Duan et al., 2018). CDMPs are derived by parameterizing the original trajectory using the hyperbolic tangent function $\tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$. We make modifications on the hyperbolic tangent space such that the joint trajectories will always evolve within the given bound. Formally, assume the joint limits are determined by \mathbf{q}_{\min} and \mathbf{q}_{\max} . The feasible joint space in terms of the exogenous variable ξ is as follows

$$\mathbf{q}(\xi) = \mathbf{q}_e \tanh(\xi) + \mathbf{q}_o, \quad (21)$$

where $\mathbf{q}_e = \text{diag}(\frac{1}{2}(\mathbf{q}_{\max} - \mathbf{q}_{\min}))$ and $\mathbf{q}_o = \frac{1}{2}(\mathbf{q}_{\max} + \mathbf{q}_{\min})$. Therefore, given the desired reference trajectory \mathbf{q}_d , its transformation into tanh-space is given as

$$\xi_d = \text{arctanh}(\mathbf{q}_e^{-1}(\mathbf{q}_d - \mathbf{q}_o)). \quad (22)$$

Consequently, Dynamic Movement Primitives (DMPs) are trained in tanh-space with modifications from potential field exerted therein. DMPs are described by

$$\begin{aligned} \tau \dot{\omega} &= -\alpha \omega, \\ \tau^2 \ddot{\xi} &= \mathbf{K}^p(\mathbf{g} - \xi) - \tau \mathbf{K}^v \dot{\xi} + \omega \mathbf{f}(\omega) + \mathbf{p}, \\ f_k(\omega) &= \frac{\sum_{i=1}^N \phi_i(\omega) \chi_{k,i}}{\sum_{i=1}^N \phi_i(\omega)}, \end{aligned} \quad (23)$$

where $\tau > 0$ denotes movement duration, $\alpha > 0$ is a scalar, ω is the phase variable on which the forcing term \mathbf{f} depends, $\phi_i(\omega) = e^{-h_i(\omega-l_i)^2}$ is the basis function with $h_i > 0$ and $l_i \in [0, 1]$ and $\chi_{k,i}$ is the corresponding weight. Moreover, \mathbf{p} is the coupling term that modifies the DMPs trajectory according to the potential field. It should be noted that the training of DMP is the same as the usual procedure, i.e., fitting the acceleration profile with Gaussian radial basis functions. Here the only difference is that it happens in tanh-space. The working flow of CDMPs is illustrated in Figure 3 and the corresponding algorithm is summarized in Algorithm 2.

Algorithm 2: CDMPs

Input: Feasible trajectories $\hat{\mathbf{q}}$

Output: Bounded trajectories $\tilde{\mathbf{q}}$

Initialization : \mathbf{q}_{\min} , \mathbf{q}_{\max} , DMPs hyperparameters

- 1: Transform $\hat{\mathbf{q}}$ into tanh-space as per (22)
 - 2: Train DMPs as per (23)
 - 3: Add coupling term for obstacle avoidance as per (12)
 - 4: Transform back for $\tilde{\mathbf{q}}$ as per (21)
 - 5: **return** $\tilde{\mathbf{q}}$
-

3.2.3. Reinforcement Learning of Hyperparameters

Normally, the strength of a potential field determined by the hyperparameters can affect the performance of obstacle avoidance greatly. If the strength is too high, then the robot will react more than enough to avoid the obstacle. Yet, although joint limit cannot be exceeded thanks to CDMPs, the similarity with respect to the demonstrated behavior will be very much sacrificed. On the other hand, if the strength is too low, the robot would be in danger of hitting the obstacle. In order to find the optimal hyperparameter values, a non-parametric reinforcement learning algorithm called Cost-regularized Kernel Regression (CrKR) is employed (Kober et al., 2012). CrKR is used to learn the optimal policy that maps the state, which is the position of the obstacle, to the optimal action, which is associated with the corresponding hyperparameters. At the beginning of each trial, for a query obstacle position \mathbf{s}_o , the hyperparameters are sampled from the Gaussian distribution $\gamma_i \sim \mathcal{N}(\gamma_i(\mathbf{s}_o), \sigma^2(\mathbf{s}_o))$ with the mean given by

$$\gamma_i(\mathbf{s}_o) = \mathbf{k}(\mathbf{s}_o)^\top (\mathbf{K} + \lambda \mathbf{C})^{-1} \mathbf{\Gamma}_i, \quad (24)$$

and the variance to facilitate exploration given by

$$\sigma^2(\mathbf{s}_o) = k(\mathbf{s}_o, \mathbf{s}_o) + \lambda - \mathbf{k}(\mathbf{s}_o)^\top (\mathbf{K} + \lambda \mathbf{C})^{-1} \mathbf{k}(\mathbf{s}_o), \quad (25)$$

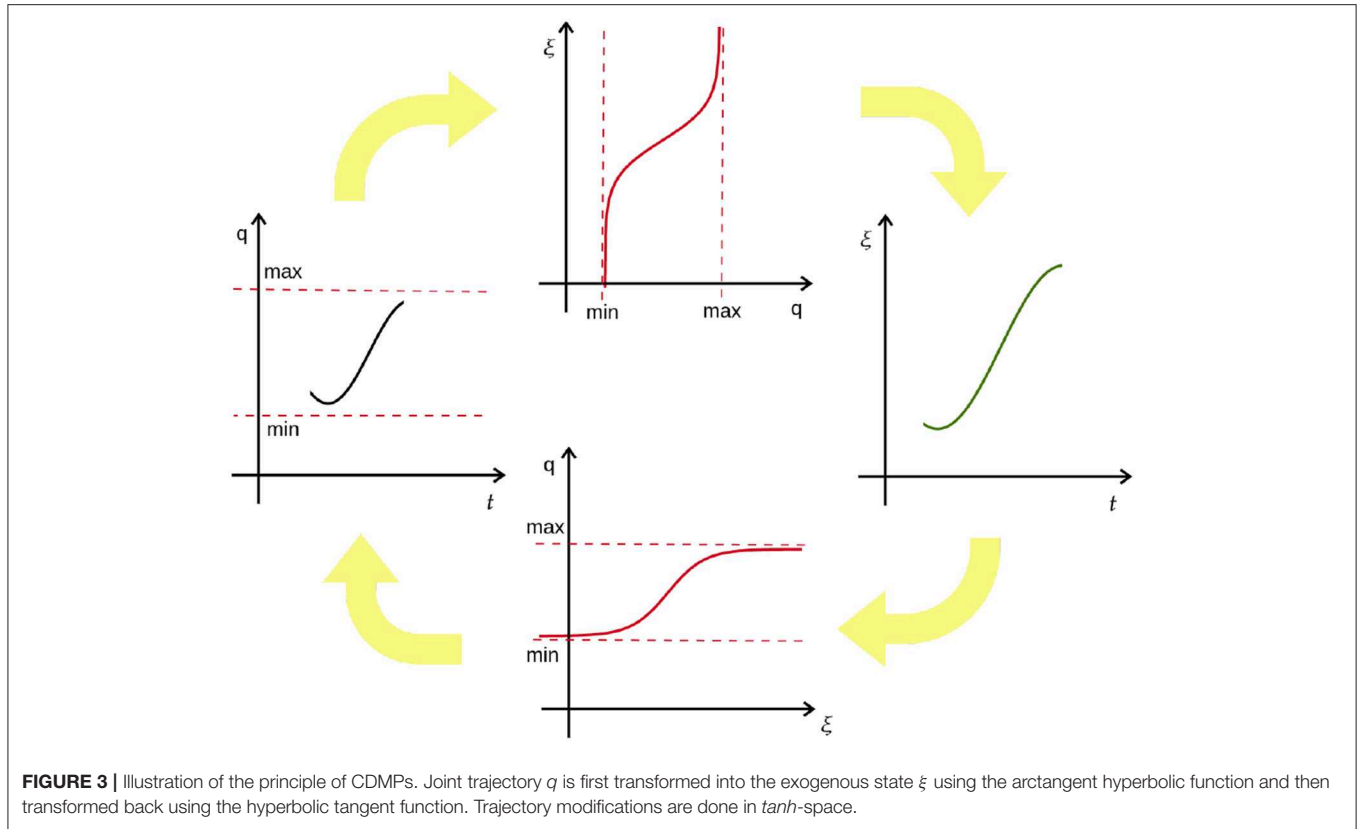


FIGURE 3 | Illustration of the principle of CDMPs. Joint trajectory q is first transformed into the exogenous state ξ using the arctangent hyperbolic function and then transformed back using the hyperbolic tangent function. Trajectory modifications are done in \tanh -space.

where $k(\cdot, \cdot)$ is a kernel function; $\mathbf{k}(s_0) = \boldsymbol{\phi}(s_0)^T \boldsymbol{\Phi}^T$ with $\boldsymbol{\phi}(\cdot)$ being basis functions and $\boldsymbol{\Phi}_i = \boldsymbol{\phi}(s_i)$, $\mathbf{K} = \boldsymbol{\Phi} \boldsymbol{\Phi}^T$, $\mathbf{C} = \text{diag}(c_1, c_2, \dots, c_n)$ is a cost matrix, λ is a ridge factor and $\boldsymbol{\Gamma}_i$ is a vector containing the training examples for the hyperparameters. It should be noted that the cost element c_i stored in the cost matrix \mathbf{C} is calculated by the Procrustes distance between the collision-free trajectory $\tilde{\tau}_i$ and the demonstrated trajectory $\hat{\tau}$, i.e., $c_i = d(\tilde{\tau}_i, \hat{\tau})$. The learning process terminates when the distance between the robot end-effector and the obstacle is smaller than the threshold.

Algorithm 3: Learning to avoid obstacles with minimal intervention control

- 1: **Trajectory retrieval:**
 - Collect dataset $\{\{t_{n,m}, \boldsymbol{\tau}_{t,m}^x, \boldsymbol{\tau}_{t,m}^j\}_{t=1}^M\}_{m=1}^M$.
 - Extract statistical information as per (8).
 - Fuse Cartesian and joint trajectories as per (11).
- 2: **Obstacle avoidance**
 - Choose potential field as per Algorithm 1.
 - Train CDMPs as per Algorithm 2.
 - Hyperparameters learning as per (24) and (25).
- 3: **Minimal intervention control:**
 - Apply control law as per (31).

3.3. Minimal Intervention Control

To track the reference trajectory $\tilde{\boldsymbol{\mu}}$ with optimal control inputs $\mathbf{U}^* = [\mathbf{u}_t^T \mathbf{u}_{t+1}^T \dots \mathbf{u}_{t+m-1}^T]^T$, a minimal intervention control strategy can be employed (Calinon et al., 2014). To start with, an optimization problem is formulated as

$$\mathbf{U}^* = \arg \min_{\mathbf{U} \in \mathbb{R}^{m \times n}} \sum_{i=t}^{t+m} ((\boldsymbol{\tau}_i^j - \tilde{\boldsymbol{\mu}}_i)^T \hat{\boldsymbol{\Sigma}}_i^{-1} (\boldsymbol{\tau}_i^j - \tilde{\boldsymbol{\mu}}_i)) + \sum_{i=t}^{t+m-1} \mathbf{u}_i^T \mathbf{R}_i \mathbf{u}_i, \tag{26}$$

where \mathbf{R}_i is a positive definite weight matrix. To find the analytical solution \mathbf{U}^* , we formulate joint dynamics in terms of a double integrator:

$$\begin{bmatrix} \mathbf{q}_{t+1} \\ \dot{\mathbf{q}}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I} & \mathbf{I} \Delta t \\ \mathbf{0} & \mathbf{I} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{q}_t \\ \dot{\mathbf{q}}_t \end{bmatrix}}_{\boldsymbol{\tau}_t^j} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{I} \Delta t \end{bmatrix}}_{\mathbf{B}} \mathbf{u}_t. \tag{27}$$

By applying (27) repetitively, we have

$$\underbrace{\begin{bmatrix} \boldsymbol{\tau}_t^j \\ \boldsymbol{\tau}_{t+1}^j \\ \boldsymbol{\tau}_{t+2}^j \\ \vdots \\ \boldsymbol{\tau}_{t+m}^j \end{bmatrix}}_{\bar{\boldsymbol{\tau}}^j} = \underbrace{\begin{bmatrix} \mathbf{I} \\ \mathbf{A} \\ \mathbf{A}^2 \\ \vdots \\ \mathbf{A}^m \end{bmatrix}}_{\bar{\mathbf{A}}} \boldsymbol{\tau}_t^j + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{AB} & \mathbf{B} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}^{m-1} \mathbf{B} & \mathbf{A}^{m-2} \mathbf{B} & \dots & \mathbf{B} \end{bmatrix}}_{\bar{\mathbf{B}}} \underbrace{\begin{bmatrix} \mathbf{u}_t \\ \mathbf{u}_{t+1} \\ \vdots \\ \mathbf{u}_{t+m-1} \end{bmatrix}}_{\mathbf{U}}. \tag{28}$$

Therefore, the original optimization problem as in (26) can be re-written compactly:

$$f(\mathbf{U}) = ((\bar{\tau}^j - \bar{\mu})^\top \bar{\Sigma}^{-1} (\bar{\tau}^j - \bar{\mu})) + \mathbf{U}^\top \bar{\mathbf{R}} \mathbf{U} \quad (29)$$

where

$$\begin{aligned} \bar{\mu} &= [\tilde{\mu}_t^\top, \tilde{\mu}_{t+1}^\top, \dots, \tilde{\mu}_{t+m}^\top]^\top, \\ \bar{\Sigma} &= \text{blockdiag}(\hat{\Sigma}_t, \hat{\Sigma}_{t+1}, \dots, \hat{\Sigma}_{t+m}), \\ \bar{\mathbf{R}} &= \text{blockdiag}(\mathbf{R}_t, \mathbf{R}_{t+1}, \dots, \mathbf{R}_{t+m-1}). \end{aligned} \quad (30)$$

By inserting (28) into (29) and setting the derivative with respect to \mathbf{U} equal to 0, the optimal control policy \mathbf{U}^* can be calculated as

$$\mathbf{U}^* = (\bar{\mathbf{B}}^\top \bar{\Sigma}^{-1} \bar{\mathbf{B}} + \bar{\mathbf{R}})^{-1} (\bar{\mathbf{B}}^\top \bar{\Sigma}^{-1} (\bar{\mu} - \bar{\mathbf{A}} \bar{\tau}_t^j)). \quad (31)$$

The complete procedure for the proposed approach to obstacle avoidance with minimal intervention control is summarized in Algorithm 3.

4. EXPERIMENTS

This section illustrates the effectiveness of the proposed approach by reporting the results of two evaluative experiments. The first experiment is a toy example where the comparison between different potential fields is conducted. With the help of the toy example, the necessity of introducing a novel imitation metric under obstacle avoidance is unveiled. Subsequently, the second experiment is devised as a transportation task under obstacle avoidance, following the proposed framework to obstacle avoidance under PbD. The real experiment is conducted on an iCub, a full-body child-size humanoid robot (Natale et al., 2017).

4.1. Toy Example

By convention, the imitation metric between the reproduced trajectory and the demonstrated one is designed as RMSE. However, in the context of obstacle avoidance, we show in the toy example that RMSE is not a proper imitation metric as it can not reflect the real imitation fidelity. Therefore, a novel imitation metric instead of RMSE is required to measure the imitation fidelity. As discussed earlier, we choose the imitation metric by resorting to the technique of curve similarity analysis. Specifically, the Procrustes distance is employed to replace the RMSE so as to compare the performance of different potential fields. In our toy example, the performances of the Fajen potential field as provided in (12) and the Khatib potential field (Khatib, 1986) are compared against each other. The mechanism of the Khatib potential field for obstacle avoidance is that the repulsive force becomes larger as the manipulator moves closer to the obstacle:

$$U(\mathbf{x}) = \begin{cases} \frac{\eta}{2} (\frac{1}{\rho(\mathbf{x})} - \frac{1}{\rho_0})^2, & \rho(\mathbf{x}) \leq \rho_0, \\ 0, & \rho(\mathbf{x}) > \rho_0, \end{cases} \quad (32)$$

where $\rho(\mathbf{x})$ is the distance between the current position and the obstacle, η is a gain term and ρ_0 is called the threshold distance

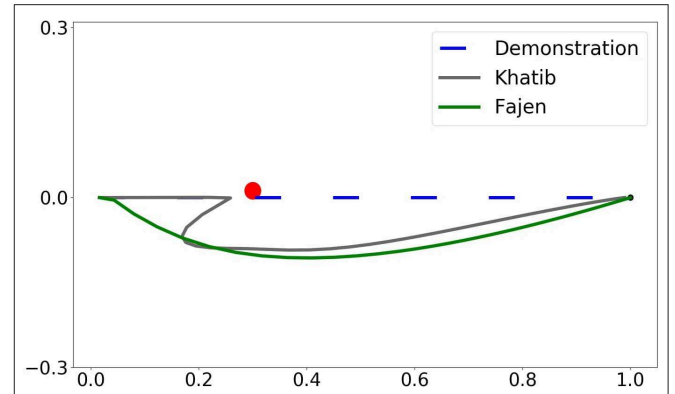


FIGURE 4 | Illustration of the comparison between Khatib and Fajen potential field with Procrustes distance.

TABLE 1 | Comparison of imitation metric.

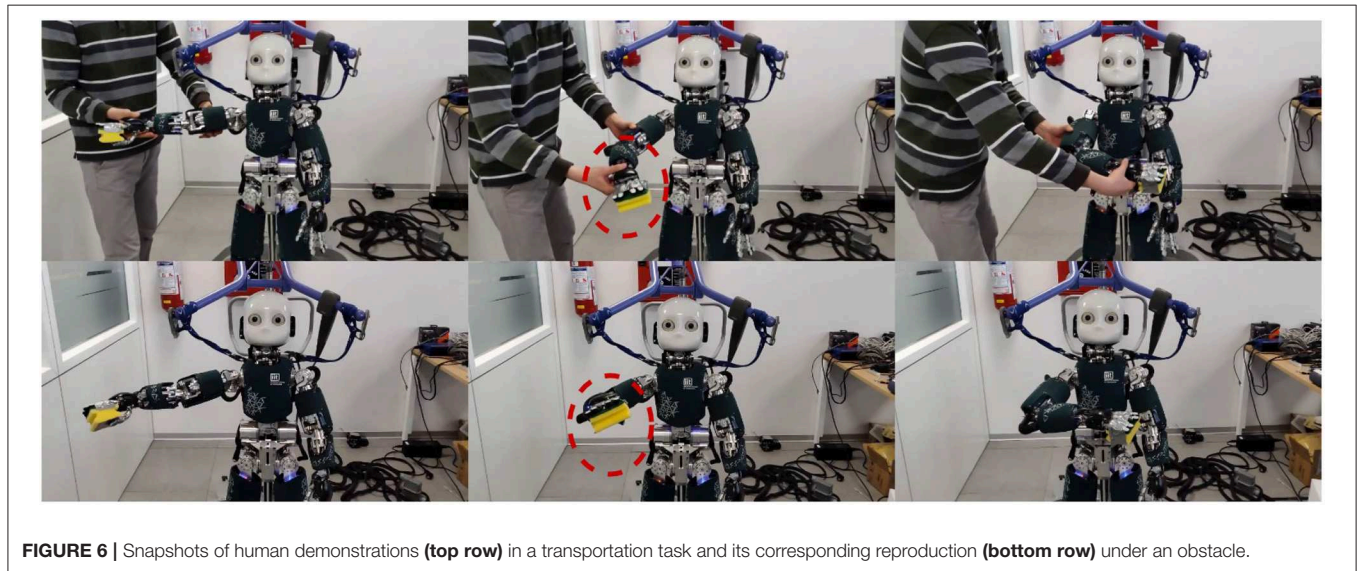
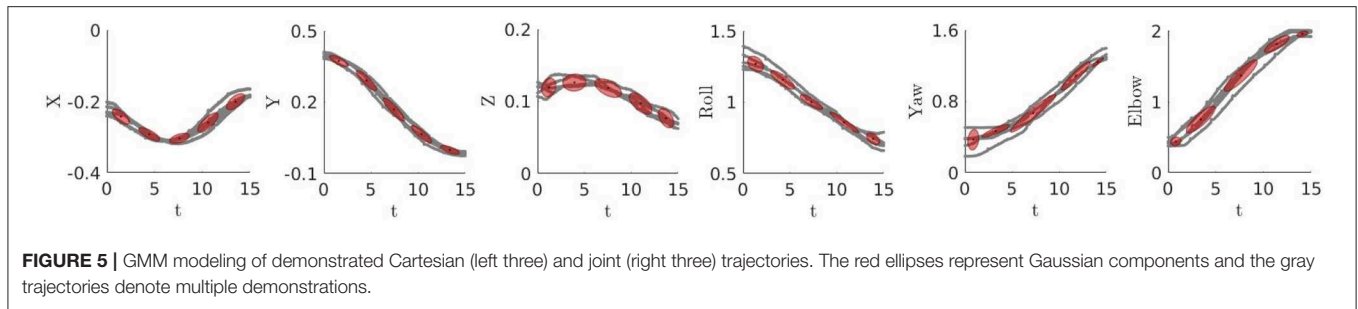
	Khatib potential field	Fajen potential field
Root mean square error	0.37	0.41
Procrustes distance	3.06	0.92

to the obstacle point. The coupling term for obstacle avoidance with Khatib potential field is calculated by deriving the gradient, i.e., $p(\mathbf{x}) = -\nabla U(\mathbf{x})$.

The comparison of obstacle avoidance performance between the Khatib potential field and the Fajen potential field is shown in Figure 4. A point moves from the starting point (0, 0) to the goal point (1, 0) with an obstacle located at midway (0.3, 0). The parameters used for Khatib potential field are chosen as $\eta = 0.12$ and $\rho_0 = 0.15$, while the parameters used for the Fajen potential field are chosen as $\gamma = 2000$ and $\beta = 20/\pi$.

As can be seen in Figure 4, by employing Khatib potential field, the area enclosed by the distorted obstacle avoidance trajectory and the demonstrated one is smaller than that of Fajen field. In this imitation metric, Khatib potential field preserves higher imitation fidelity than Fajen field. However, this metric is not reasonable as one can identify intuitively that the obstacle avoidance trajectory under Fajen field should share more similarity to the demonstrated one than Khatib field. And by employing Procrustes distance for the imitation metric, indeed the obstacle avoidance trajectory under Fajen field has a smaller deviation from the demonstrated one.

Numerically, when evaluated with root mean square error, the distance is 0.37 for Khatib potential field and 0.41 for the Fajen field. However, by using the Procrustes distance as imitation metric, the distance becomes 3.06 for the Khatib field and only 0.92 for the Fajen field. In conclusion, Procrustes distance should be chosen over RMSE when the reproduction trajectory is modified irregularly. Besides, Fajen field should be preferred for obstacle avoidance as it preserves imitation fidelity better. Table 1 summarizes the quantitative comparison between two potential fields under two types of similarity metric. The code is available as **Supplementary Material**.



4.2. Transportation Task

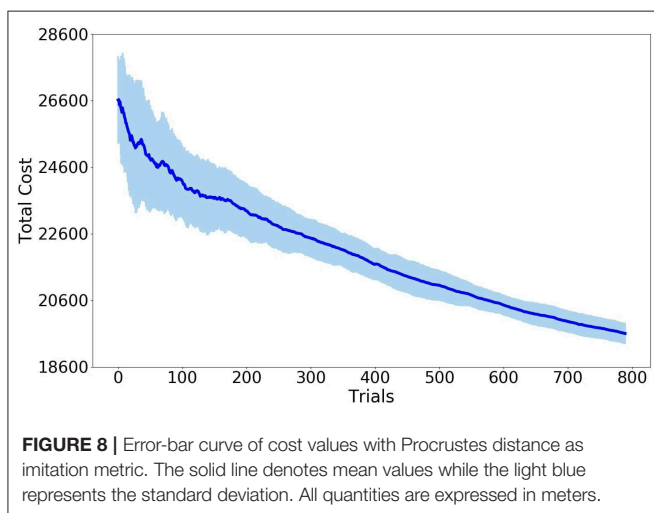
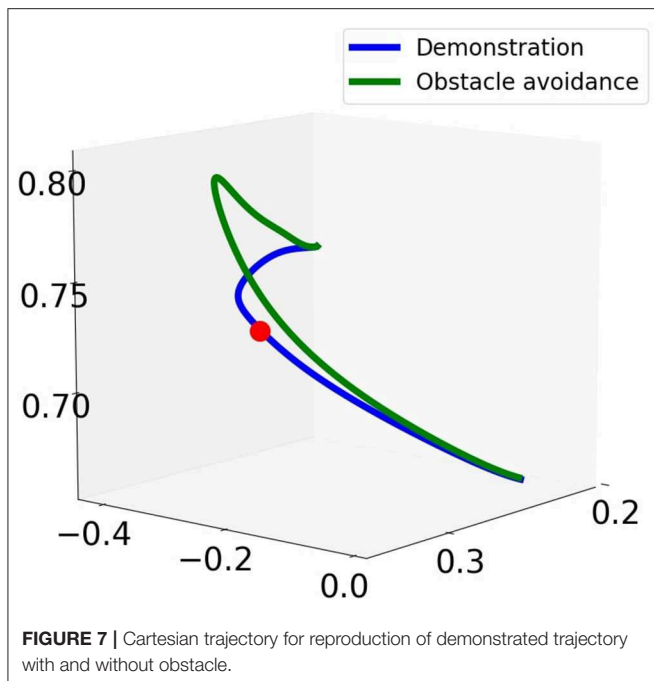
We evaluate the proposed method with a real experiment on the iCub humanoid robot. In order to show the effectiveness of the proposed framework for obstacle avoidance within PbD, a transportation task is considered as a concept-proof experiment. As iCub is a full-body humanoid robot with 56 DoFs in total, to accomplish the transportation task in our case, only the right arm having 7 DoF is activated with 3 from the shoulder, 3 from the wrist, and 1 from the elbow.

The conceived experimental set-up is as such: a sponge is first handed over to the robot, and then a human teacher guides the activated robot right arm to reach a final location. Finally, the robot is required to reproduce the demonstrated trajectory with the existence of an obstacle positioned midway through the demonstrated trajectory. During the kinesthetic teaching phase, the robot is taught the transportation task for five times. The collected dataset records both robot Cartesian and joint trajectories. In order to encode the probabilistic information underlying the multiple demonstrated trajectories, a GMM with five components is employed to model the distributions of the demonstrated Cartesian and joint trajectories, respectively. The GMM modeling results are plotted in **Figure 5**, and it can be observed that the trajectory segment with larger variation incurs larger covariance. The probabilistic reference trajectories to control the robot are extracted with GMR. In order to unify the inconsistency between Cartesian and joint trajectories as a result

of multiple demonstrations, the trajectories from both spaces are fused by the Gaussian product as in (11). The experimental illustration is shown in **Figure 6**.

In order to incorporate joint limit avoidance during obstacle avoidance, the retrieved trajectories are then used to train CDMPs according to Algorithm 2. For the design of the coupling term of CDMPs, the Fajen potential field is employed for obstacle avoidance. As it has been shown in the toy example, it outperforms Khatib potential field with the Procrustes distance as imitation metric. During the reproduction phase, a virtual obstacle is positioned at $(0.31, -0.25, 0.73)$ m in Cartesian space with respect to the world frame. The snapshots of the robot reproducing the transportation tasks with the obstacle is shown in the bottom row of **Figure 4**. The Cartesian trajectory of the obstacle avoidance trajectory (in green) and the demonstrated trajectory (in blue) is shown in **Figure 7**. During the execution of the obstacle avoidance, no joint violates the corresponding joint limits.

The hyperparameter values of the potential field are determined by the reinforcement learning algorithm CrKR in an off-line fashion. We learn these hyperparameters as a function of the position of the obstacle. The reward used by CrKR is formulated as the Procrustes distance between the obstacle avoidance trajectory and the demonstrated one. To run CrKR, we choose the Gaussian kernel $k = \exp((s_i - s_j)^2)$ for the



algorithm with the distance of a point from itself $k(\mathbf{s}, \mathbf{s}) = 1$ and a ridge factor $\lambda = 0.5$. During the preparation steps of the algorithm, the corresponding matrices \mathbf{K} , \mathbf{C} and $\mathbf{\Gamma}$ are initialized by 20 samples. The total number of trials for each run of the algorithm is 800. The cost variance is calculated by repeating the learning process for five times. The learning results are reported in **Figure 8**. The learning process terminates when the threshold of the minimum distance to the obstacle is triggered. When the learning process finishes, the Fajen field parameters are obtained as $\gamma = 1260$ and $\beta = 3.2$ with respect to the specified obstacle position.

Finally, the reproduction trajectory is executed with a minimal intervention controller in order to endow the robot with variable impedance skills. The cost function of the minimal intervention controller is parameterized by $\mathbf{R} = 10^{-2}\mathbf{I}$. The prediction horizon is set as 10 discrete time steps.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we presented an approach to obstacle avoidance in the context of hybrid imitation learning. To exploit the probabilistic information underlying the trajectories, multiple demonstrations are taught to the robot. The initial trajectory is then retrieved from the human demonstration dataset by fusing Cartesian and joint constraints with the Gaussian product. Since there are various types of potential field, the Procrustes distance other than RMSE is employed for the benchmarking of the performance of different potential fields. As a common technique in curve similarity analysis, the Procrustes distance can better reflect the imitation fidelity between the obstacle-avoidance trajectory and the original demonstrated one. It should be noted that minimizing the Procrustean imitation metric might be more numerically expensive than square root distance. Given that the potential field would modify joint trajectories unpredictably during obstacle avoidance, joint limit avoidance is incorporated to guarantee the evolution of the modified trajectories is always bounded within the allowable range. To this end, the novel Constrained Dynamic Movement Primitives (CDMPs) method is employed. CDMPs parameterize joint trajectories with the help of the hyperbolic tangent function. By exploiting the boundedness property of the hyperbolic tangent function, the modified joint trajectories are guaranteed to evolve within the specified range. Further, in view of the fact that the performance of obstacle avoidance is quite susceptible to the hyperparameters of the potential field, a reinforcement learning algorithm is used to find the most suitable hyperparameters. The final obstacle avoidance trajectory is tracked with a minimal intervention controller to endow the robot with variable impedance capabilities.

As a preliminary attempt to address obstacle avoidance issues under PbD, a number of topics remain to be investigated for future work. For example, the position of the obstacle is given beforehand in this paper, but it could be interesting to exploit the visual system of iCub such that the obstacle position can be autonomously determined. In addition, the proposed method could be considered for extension to the case of moving obstacles or obstacles with 3D shape.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

Part of this work has been carried out at the Machine Learning Genoa (MaLGa) Center, Università di Genova (IT). LR acknowledges the financial support of the European Research Council (grant SLING 819789), the AFOSR projects

FA9550-17-1-0390 and BAA-AFRL-AFOSR-2016-0007 (European Office of Aerospace Research and Development), the EU H2020-MSCA-RISE project NoMADS - DLV-777826, and the Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frobt.2020.00060/full#supplementary-material>

REFERENCES

- Abu-Dakka, F. J., Nemeč, B., Kramberger, A., Buch, A. G., Krüger, N., and Ude, A. (2014). Solving peg-in-hole tasks by human demonstration and exception strategies. *Ind. Robot Int J.* 41, 575–584. doi: 10.1108/IR-07-2014-0363
- Abu-Dakka, F. J., Rozo, L., and Caldwell, D. G. (2018). Force-based variable impedance learning for robotic manipulation. *Robot. Auton. Syst.* 109, 156–167. doi: 10.1016/j.robot.2018.07.008
- Billard, A., Calinon, S., Dillmann, R., and Schaal, S. (2008). “Robot programming by demonstration,” in *Springer Handbook of Robotics*, eds B. Siciliano and O. Khatib (Berlin; Heidelberg: Springer), 1371–1394. doi: 10.1007/978-3-540-30301-5_60
- Calinon, S., and Billard, A. (2008). “A probabilistic programming by demonstration framework handling constraints in joint space and task space,” in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Nice: IEEE), 367–372. doi: 10.1109/IROS.2008.4650593
- Calinon, S., Bruno, D., and Caldwell, D. G. (2014). “A task-parameterized probabilistic model with minimal intervention control,” in *2014 IEEE International Conference on Robotics and Automation (ICRA)* (Hong Kong: IEEE), 3339–3344. doi: 10.1109/ICRA.2014.6907339
- Calinon, S., and Lee, D. (2016). “Learning control,” in *Humanoid Robotics: A Reference*, eds A. Goswami and P. Vadakkepat (Dordrecht: Springer Netherlands), 1–52. doi: 10.1007/978-94-007-7194-9_68-1
- Dryden, I. L. (2014). *Shape Analysis*. Wiley StatsRef: Statistics Reference Online. doi: 10.1002/9781118445112.stat05087
- Duan, A., Camoriano, R., Ferigo, D., Calandriello, D., Rosasco, L., and Pucci, D. (2018). “Constrained DMPs for feasible skill learning on humanoid robots,” in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)* (Beijing: IEEE), 1–6. doi: 10.1109/HUMANOIDS.2018.8624934
- Duan, A., Camoriano, R., Ferigo, D., Huang, Y., Calandriello, D., Rosasco, L., et al. (2019). “Learning to sequence multiple tasks with competing constraints,” in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Macau: IEEE). doi: 10.1109/IROS40897.2019.8968496
- Fajen, B. R., and Warren, W. H. (2003). Behavioral dynamics of steering, obstacle avoidance, and route selection. *J. Exp. Psychol. Hum. Percept. Perform.* 29:343. doi: 10.1037/0096-1523.29.2.343
- Hoffmann, H., Pastor, P., Park, D.-H., and Schaal, S. (2009). “Biologically-inspired dynamical systems for movement generation: automatic real-time goal adaptation and obstacle avoidance,” in *2009 IEEE International Conference on Robotics and Automation* (Kobe: IEEE), 2587–2592. doi: 10.1109/ROBOT.2009.5152423
- Huang, Y., Büchler, D., Koç, O., Schölkopf, B., and Peters, J. (2016). “Jointly learning trajectory generation and hitting point prediction in robot table tennis,” in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)* (Cancun: IEEE), 650–655. doi: 10.1109/HUMANOIDS.2016.7803343
- Huang, Y., Rozo, L., Silvério, J., and Caldwell, D. G. (2019). Kernelized movement primitives. *Int. J. Robot. Res.* 38, 833–852. doi: 10.1177/0278364919846363
- Huang, Y., Silvério, J., and Caldwell, D. G. (2018). “Towards minimal intervention control with competing constraints,” in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid: IEEE), 733–738. doi: 10.1109/IROS.2018.8594235
- Huber, L., Billard, A., and Slotine, J.-J. E. (2019). Avoidance of convex and concave obstacles with convergence ensured through contraction. *IEEE Robot. Autom. Lett.* 4, 1462–1469. doi: 10.1109/LRA.2019.2893676
- Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., and Schaal, S. (2013). Dynamical movement primitives: learning attractor models for motor behaviors. *Neural Comput.* 25, 328–373. doi: 10.1162/NECO_a_00393
- Kalakrishnan, M., Chitta, S., Theodorou, E., Pastor, P., and Schaal, S. (2011). “Stomp: stochastic trajectory optimization for motion planning,” in *2011 IEEE International Conference on Robotics and Automation (IEEE)*, 4569–4574. doi: 10.1109/ICRA.2011.5980280
- Kavraki, L. E., Svestka, P., Latombe, J., and Overmars, M. H. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Trans. Robot. Autom.* 12, 566–580. doi: 10.1109/70.508439
- Khansari-Zadeh, S. M., and Billard, A. (2012). A dynamical system approach to realtime obstacle avoidance. *Auton. Robots* 32, 433–454. doi: 10.1007/s10514-012-9287-y
- Khatib, O. (1986). “Real-time obstacle avoidance for manipulators and mobile robots,” in *Autonomous Robot Vehicles* (New York, NY: Springer), 396–404. doi: 10.1007/978-1-4613-8997-2_29
- Kim, J.-O., and Khosla, P. K. (1992). Real-time obstacle avoidance using harmonic potential functions. *IEEE Trans. Robot. Autom.* 8, 338–349. doi: 10.1109/70.143352
- Kober, J., Wilhelm, A., Oztop, E., and Peters, J. (2012). Reinforcement learning to adjust parametrized motor primitives to new situations. *Auton. Robots* 33, 361–379. doi: 10.1007/s10514-012-9290-3
- Koren, Y., and Borenstein, J. (1991). “Potential field methods and their inherent limitations for mobile robot navigation,” in *Proceedings 1991 IEEE International Conference on Robotics and Automation* (Sacramento, CA: IEEE), 1398–1404. doi: 10.1109/ROBOT.1991.131810
- LaValle, S. M. (1998). *Rapidly-Exploring Random Trees: A New Tool for Path Planning*. Department of Computer Science, Iowa State University.
- Mitchel, T. W., Ruan, S., and Chirikjian, G. S. (2018). “Signal alignment for humanoid skeletons via the globally optimal reparameterization algorithm,” in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)* (Beijing: IEEE), 217–223. doi: 10.1109/HUMANOIDS.2018.8624999
- Natale, L., Bartolozzi, C., Pucci, D., Wykowska, A., and Metta, G. (2017). iCub: The not-yet-finished story of building a robot child. *Sci. Robot.* 2:eaaq1026. doi: 10.1126/scirobotics.aaq1026
- Rai, A., Meier, F., Ijspeert, A., and Schaal, S. (2014). “Learning coupling terms for obstacle avoidance,” in *2014 IEEE-RAS International Conference on Humanoid Robots* (Atlanta, GA: IEEE), 512–518. doi: 10.1109/HUMANOIDS.2014.7041410
- Rasmussen, C. E. (2003). “Gaussian processes in machine learning,” in *Summer School on Machine Learning* (Berlin; Heidelberg: Springer), 63–71. doi: 10.1007/978-3-540-28650-9_4
- Ruan, S., Chirikjian, G. S., and Ding, J. (2018). “Lower bounds of the allowable motions of one n-dimensional ellipsoid contained in another,” in *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Quebec City, QC: American Society of Mechanical Engineers), V05BT07A081-V05BT07A081. doi: 10.1115/DETC2018-85851
- Schulman, J., Duan, Y., Ho, J., Lee, A., Awval, I., Bradlow, H., et al. (2014). Motion planning with sequential convex optimization and convex collision checking. *Int. J. Robot. Res.* 33, 1251–1270. doi: 10.1177/0278364914528132

Zucker, M., Ratliff, N., Dragan, A. D., Pivtoraiko, M., Klingensmith, M., Dellin, C. M., et al. (2013). Chomp: covariant hamiltonian optimization for motion planning. *Int. J. Robot. Res.* 32, 1164–1193. doi: 10.1177/0278364913488805

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Duan, Camoriano, Ferigo, Huang, Calandriello, Rosasco and Pucci. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.