

# Learning Reactive Robot Behavior for Autonomous Valve Turning

Seyed Reza Ahmadzadeh, Petar Kormushev, Rodrigo S. Jamisola and Darwin G. Caldwell

**Abstract**—A learning approach is proposed for the challenging task of autonomous robotic valve turning in the presence of active disturbances and uncertainties. The valve turning task comprises two phases: reaching and turning. For the reaching phase the manipulator learns how to generate trajectories to reach or retract from the target. The learning is based on a set of trajectories demonstrated in advance by the operator. The turning phase is accomplished using a hybrid force/motion control strategy. Furthermore, a reactive decision making system is devised to react to the disturbances and uncertainties arising during the valve turning process. The reactive controller monitors the changes in force, movement of the arm with respect to the valve, and changes in the distance to the target. Observing the uncertainties, the reactive system modulates the valve turning task by changing the direction and rate of the movement. A real-world experiment with a robot manipulator mounted on a movable base is conducted to show the efficiency and validity of the proposed approach.

## I. INTRODUCTION

Robotic valve turning is a challenging task specially in unstructured environments with increasing level of uncertainty (e.g., underwater). The existing disturbances in the environment or the noise in the sensors can endanger both the robot and the valve during the operation. For instance, the vision system may be occluded and thus introduce a delay in updating the data, or even providing the system with wrong information. Exerting huge forces/torques on the valve by the robot, is another hazardous and highly probable situation. In such cases an autonomous system that is capable of observing the current state of the system and reacting accordingly, can help to accomplish the mission successfully even in the presence of noise.

Robotic valve manipulation contains a number of complex and challenging subtasks. There seem to be few published description of attempts directly related to this task. Prior works in industrial robotic valve operation, generally use nonadaptive classical control and basic trajectory planning methods. In [2], Abidi *et al.*, tried to achieve inspection and manipulation capabilities in the semi-autonomous operation of a control panel in a nuclear power plant. A 6-DoF industrial robot equipped with a number of sensors (e. g., vision, range, sound, proximity, force/torque, and touch) was used. The main drawback is that their approach is developed for static environments with predefined dimensions and scales. For instance, the size and position of the panel, the valve,

This research was partially supported by the PANDORA EU FP7 project [1] under the grant agreement No. ICT-288273. <http://persistentautonomy.com/>

The authors are with the Department of Advanced Robotics, Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genova, Italy. {reza.ahmadzadeh, rodrigo.jamisola, petar.kormushev, darwin.caldwell} @iit.it

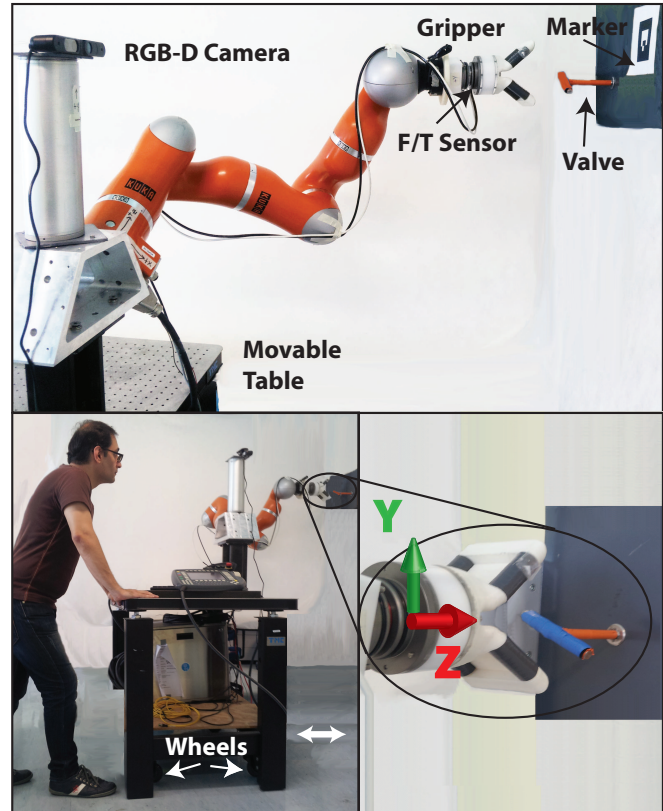


Fig. 1: The experimental set-up for the valve turning task. The valve is detected and localized using an RGB-D sensor through an AR-marker. The manipulator is equipped with a gripper and is mounted on a movable (wheeled) table. During the execution of the task, a human can create a random disturbance by perturbing the base of the robot.

and other objects in the room are manually engineered into the system. More recent approaches generally use sensor-based movement methods which implies that the robot trajectories have not been programmed off-line. In [3], the robot is equipped with a torque sensor and the valve which is equipped with a proximity sensor is detected using a vision sensor. The authors focus on a model-based approach to avoid over-tightening/loosening of the valve. The other phases of the valve manipulation process are accomplished using classical methods. In their next research [4] the authors develop the valve manipulation task in outdoor environment. The vision sensor is replaced with a thermal camera, and the (round) valve is replaced with a T-bar valve, which is easier for the robot to manipulate. The main focus of [4] is detecting the valve and avoiding the over-tightening/loosening of the

valve in an early stage using a model-based technique.

Other groups have also investigated valve turning. In [5] a framework for valve turning is proposed using a dual-arm aerial manipulator system. The framework is built based on teleoperation and employs motion detection, voice control and joystick inputs. A user-guided manipulation framework is proposed in [6]. Although the planning algorithm generates the robot motions autonomously, the search process and the object detection phase are accomplished by a human operator and the result is passed to the robot. A dual-arm impedance hierarchical controller is devised in [7] that employs the upper body kinematics and dynamics of a humanoid robot for reaching and turning a valve.

Compared to our previous research [8], [9], this work provides the following three contributions: (i) in our previous research the turning phase was done by programming the turning motion into the robot. In this paper, on the other hand, a force control strategy is proposed for handling the turning phase; (ii) similar to the previous research a Reactive Fuzzy Decision Maker (RFDM) system is designed in order to react to external disturbances and sudden movements. The reactive system in the previous work monitors only the relative movement between the gripper and the valve, whereas, the new reactive system, also takes into consideration the distance between the gripper and the valve, and the exerted forces to the end-effector besides the relative movement. The acquired RFDM system shows more efficiency and better sensitivity which results in a safer valve turning process; (iii) in our previous research an Optitrack system was used which captures real-time 3D position and orientation of a rigid body using a number of motion capture cameras and a set of markers. Although the Optitrack system is very precise, it cannot be used in outdoor environment. In this work, on the other hand, an RGB-D sensor is used that provides the experiment with a more realistic conditions.

All presented experiments in this work were conducted in a lab environment. The experimental set-up for all the experiments is shown in Figure 1. Our future work includes investigating and accomplishing the autonomous robotic valve manipulation in underwater environment which is one of the most challenging tasks defined in the PANDORA [1] project.

## II. METHODOLOGY

The valve turning task comprises two main phases: reaching and turning. First, the robot have to learn how to reach the valve. Imitation learning approach which is designed specially to learn trajectory-based tasks, is a promising choice to learn the reaching skill [10]. In order to reproduce the reaching skill towards the target, the robot utilizes feedback from the RGB-D sensor.

When the robot is able to reproduce the reaching skill a hybrid force/motion control strategy handles the turning phase. Hybrid force/motion control is a well-established method [11], [12], [13]. Using such hybrid strategy, the force controller can maintain the contact between the valve and the gripper while the motion controller turns the valve.

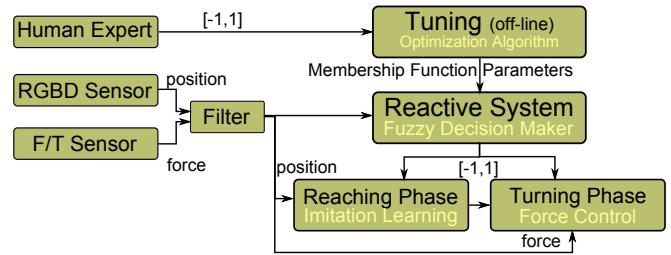


Fig. 2: A high-level flow diagram illustrating the different components of the proposed approach.

The hybrid force/motion controller utilizes feedback from a Force/Torque (F/T) sensor mounted between the end-effector and the gripper. Subsequent to the turning phase, the robot employs the reaching skill in reverse to retract from the valve.

In order to develop an autonomous system, the robot needs to deal with uncertainties. To emulate the uncertainties in our experiments, we manually apply disturbances to the system. The disturbances during the execution of the task are monitored and handled by a Reactive Fuzzy Decision Maker (RFDM). Although such reactive system can be implemented using a thresholding method, the fuzzy system is chosen. The reason is that the fuzzy system provides a continuous decision surface and it infers from a set of human-defined linguistic rules. The RFDM module, monitors the position of the gripper and the valve together with the magnitude of the forces and torques applied to the end-effector from the valve. Using this information, RFDM generates decisions that regulate the movements of the robot during the process. For example, RFDM halts the process when the magnitude of the force increases due to an undesired movement. In addition, RFDM also controls the rate of the motion. For instance, when there is no external disturbance, the robot can reach the valve faster.

As depicted in Figure 1 the experimental set-up for all the conducted experiments consists of a 7-DoF KUKA-LWR manipulator mounted on a movable (wheeled) table, a (T-bar shaped) mock-up valve mounted on the wall in the robot's workspace, a gripper designed for grasping and turning the valve, an ATI Mini45 Force/Torque (F/T) sensor which is sandwiched between the gripper and the robots end-effector, and an ASUS Xtion RGB-D sensor for detecting and localizing the valve.

Figure 2 illustrates a flow diagram of the proposed approach. The RGB-D sensor detects the pose of the valve which is used by the reaching module and RFDM. The F/T sensor monitors the force/torque applied to the gripper, which is used by the turning module and RFDM. Observing the inputs provided by the sensors, RFDM generates proper decisions in order to modulate the behavior of the robot during the process. The RFDM system is tuned by collecting data from a human expert using optimization techniques.

## III. IMITATION LEARNING

Imitation learning enables manipulators to learn and reproduce trajectory-based skills from a set of demonstra-

tions [10]. The demonstrations are provided either by teleoperation or through kinesthetic teaching. One of the most widely-used representations for trajectory-based skills is Dynamical Movement Primitives (DMP) [14]. DMP allows to learn a compact representation of the reaching skill using the recorded demonstrations. In this paper, we use the extended DMP approach proposed in [15] which also encapsulates variation and correlation information of the demonstrated skill as a mixture of dynamical systems. In order to reach a target, in this approach a set of virtual attractors is utilized. The influence of these attractors is smoothly switched along the movement on a time basis. A proportional-derivative controller is used to move the end-effector towards the target. In contrast to the original DMP, a full stiffness matrix associated with each primitives is considered. This allows to capture the variability and correlation information along the movement. The set of attractors is learned through weighted least-square regression, by using the residual errors as covariance information to estimate stiffness gain matrices.

During the demonstration phase, multiple desired trajectories are demonstrated by a human operator through kinesthetic teaching. Each demonstration  $m \in \{1, \dots, M\}$  consists of a set of  $T_m$  positions  $x$ , velocities  $\dot{x}$ , and accelerations  $\ddot{x}$ , of the end-effector in Cartesian space where  $x \in \mathbb{R}^3$ . A dataset is formed by concatenating the  $P = \sum_{m=1}^M T_m$  data points. A desired acceleration is computed based on a mixture of  $L$  proportional-derivative systems as follows:

$$\hat{\ddot{x}} = \sum_{i=1}^L h_i(t) [K_i^P (\mu_i^x - x) - k^v \dot{x}], \quad (1)$$

where  $\hat{\ddot{x}}$  is the desired acceleration,  $K_i^P$  are the stiffness matrices,  $\mu_i^x$  are the centers of the attractors in Cartesian space,  $h_i(t)$  are the weighting functions, and  $k^v$  is the derivative gain.

During the demonstration phase, the trajectories are recorded independent of the explicit time. Instead, in order to create an implicit time-dependency,  $t = -\ln(s)/\alpha$ , a canonical system is defined as follows:

$$\dot{s} = -\alpha s \quad (2)$$

where  $s$  is the decay term initialized by  $s = 1$  that monotonically converges to 0. Furthermore, a set of Gaussian basis functions are defined as  $\mathcal{N}(\mu_i^T, \Sigma_i^T)$  in time space, where the centers  $\mu_i^T$  are equally distributed in time and the variance parameters  $\Sigma_i^T$  are set to a constant value inversely proportional to the number of states.  $\alpha$  is a fixed value which depends on the duration of the demonstrations.

By determining the weighting functions  $h_i(t)$  through the decay term  $s$ , the system sequentially converges to the set of attractors. Stiffness matrices  $K_i^P$  and the centers  $\mu_i^x$  are learned from the observed data using weighted least-square regression. In the reproduction phase the system uses the learned weights and set of attractors to reproduce a trajectory to reach the target.

The recorded set of demonstrations is depicted as black curves in Figure 3. Following the described approach, the

system learns a set of attractors which can be seen in the 2D plots in Figures 3a and 4a as blue ellipsoids. Using the learned set of attractors the robot is able to reproduce a new trajectory from an arbitrary initial position towards the target. Each red trajectory in Figures 3 and 4 illustrates a reproduction. In both figures the goal, (the valve in our experiments), is shown in yellow. A snapshot of the reaching skill reproduced by the robot is shown in Figure 5.

In this paper, we introduce a new capability based on the implicit timing: a reversible behavior of the system. This new capability enables the robot to perform the following: (i) reactive behavior by switching the direction of the movement towards the target or away from it; (ii) after the task is finished the robot uses this capability to retract the arm. The advantage of the presented capability is that by learning just the reaching skill the robot is capable of reproducing multiple behaviors including reaching and retracting, and switching between them. This can be achieved by changing the timing equation from  $t = -\ln(s)/\alpha$  to  $t = t_{final} + \ln(s)/\alpha$ .

Figure 4 illustrates a reproduction from an arbitrary initial position towards the target. It can be seen that, in the middle of the movement the robot reverses the motion and moves backwards. It has to be commented that, by executing the reverse motion, the robot goes back to the center of the first attractor.

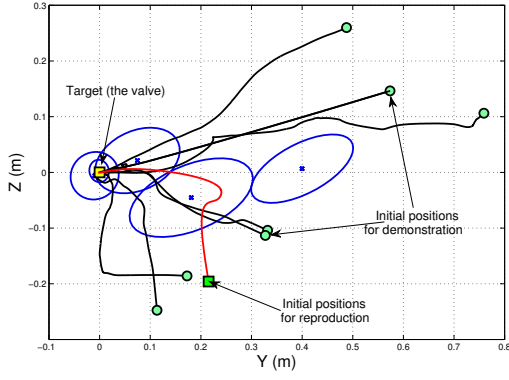
#### IV. FORCE/MOTION CONTROL STRATEGY

Once the robot learns the reaching skill, the turning phase begins. In this phase, the goal of the robot is to turn the valve (by 180° from its initial configuration) while maintaining the position of the gripper. To control the forces and torques applied to the end-effector, a hybrid force/motion control approach is used [11], [12], [13]. Hybrid force/motion controller is preferred to be used in this application because during the turning phase a zero force controller can reduce the undesired forces and torques.

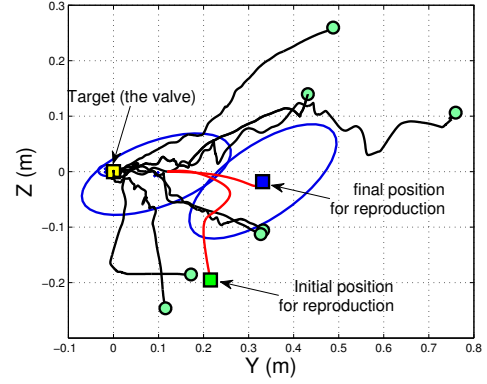
The proposed hybrid strategy is designed for 6-axes full space control. Forces and torques are controlled in the 5-axes while motion is controlled around the  $z$ -axis in order to turn the valve. The assigned coordinate system is depicted in Figure 1 which is set with respect to the initial pose of the gripper. The  $z$ -axis (surge and roll) is normal to the end-effector's palm. The  $y$ -axis (sway and pitch) is perpendicular to the  $z$ -axis and pointing sideways. And the  $x$ -axis (heave and yaw) is perpendicular to the  $z-y$  plane [16]. A desired normal force is set along the  $z$ -axis in order to maintain the gripper in contact with the valve. Zero forces and torques are specified along the  $x$ - and  $y$ -axes. The zero desired values of the forces and torques are designed to lessen the reactionary forces and torques along (around) the axes during the valve turning process.

The hybrid force/motion controller is suitable for autonomous mobile manipulation [17]. In underwater environment the valve turning task is more difficult due to the highly unstructured and uncertain environment. Also, the valve can be rusty and sensitive to high forces/torques.

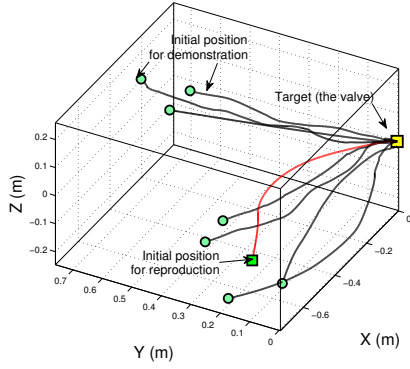
We specify the forces and torques as follows:



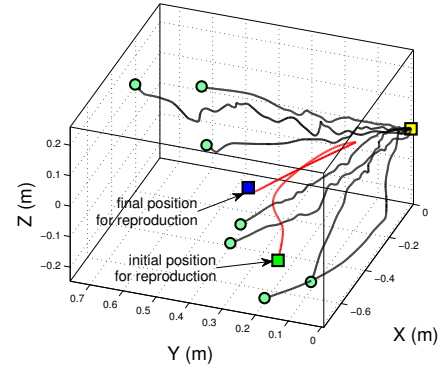
(a) Trajectories and learned attractors in 2D plane.



(a) Trajectories and learned attractors in 2D plane.



(b) Trajectories in 3D space.



(b) Trajectories in 3D space.

Fig. 3: The recorded trajectories that form the set of demonstrations (black), and the reproduced trajectory from an arbitrary initial position (red) towards the target are illustrated. The blue ellipses show the attractors and the yellow square shows the target (the valve).

Fig. 4: The recorded trajectories that form the set of demonstrations (black), and the reproduced trajectory from an arbitrary initial position (red) are illustrated. The robot retracts from the middle of the path by receiving a command from RFDM.

$$\begin{aligned} \mathbf{F}_{con} &= \mathbf{F}_{des} + \mathbf{k}_p(\mathbf{F}_{des} - \mathbf{F}_{act}) \\ \mathbf{T}_{con} &= \mathbf{T}_{des} + \mathbf{k}_p(\mathbf{T}_{des} - \mathbf{T}_{act}) \end{aligned} \quad (3)$$

where  $\mathbf{F}$  and  $\mathbf{T}$  denote forces and torques respectively, and subscripts *des*, *act*, and *con* denote the desired, actual, and control parameters respectively. In the following sections, some experimental results are explained.

#### A. Valve Turning with Stationary Base

In the first set of experiments the base of the robot remained stationary during the valve turning task. So no external disturbances are applied to the valve or the gripper. In this case, to ensure proper contact between the valve and the gripper during the turning phase a 20 N force along the z-axis is applied,  $F_{des}^z = 20$  N. The other desired forces and torques are set to zero  $F_{des}^x = F_{des}^y = T_{des}^x = T_{des}^y = 0$ . The desired roll angle is specified to be 180° clockwise with respect to the current orientation of the end-effector. The duration of turning is set to 10s. The joint angle displacements through the valve turning execution are shown in Figure 6, where the last joint moved from  $-90^\circ$  to  $90^\circ$

angle, while the rest of the joint angles move slightly in order to dissipate the forces generated at the end-effector during the execution of the task.

It can be seen in the graphs that the first 5s of the operation is used for the gripper to make contact with the valve and then dissipate the impact force. From 5s to 15s, the valve is turned by the motion controller and the generated forces and torques at the end-effector are recorded. Figure 7 (top figure) shows that the force along the z-axis is maintained at +20 N right after the impact forces have been dissipated. The controller can maintain the magnitude of the force before, during, and after the valve turning successfully.

The exertion of the necessary torque to turn the valve and the required force to maintain the contact between the valve and the gripper, generates residual forces and torques at the end-effector. In order to dissipate these reaction forces along the x- and y-axes, a set of forces in range  $-5$  to  $+5$  N are generated. During the turning phase until the 20s of the task, the z-axis torque averaged at around +1 Nm. This is shown in Figure 8 (top figure).

On the other hand, the torque around x-axis is maintained



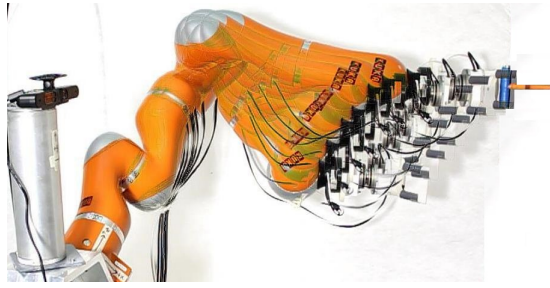


Fig. 5: The robot reaching the valve during the reproduction phase.

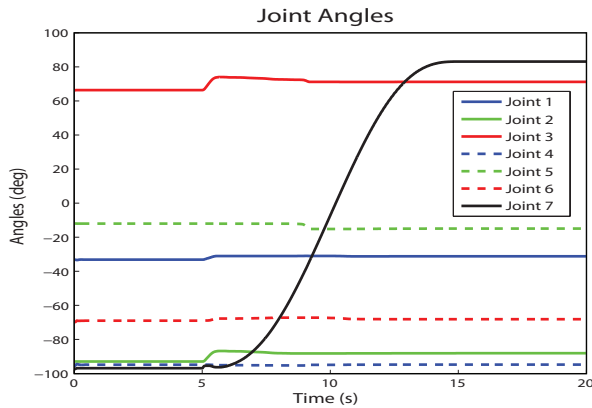


Fig. 6: Joint angles during the valve turning task with stationary base.

at  $-0.5 \text{ Nm}$  at the start of the contact until the end of the turning phase, and then decreased to  $0 \text{ Nm}$  by the end of the task. The torque around the  $y$ -axis moved from its initial torque, (after impact that is around  $-0.5 \text{ Nm}$ ) to  $-1 \text{ Nm}$  in an attempt to dissipate the residual torques. However, due to the physical limitation of the valve, (i.e., the gripper has a slot along the  $y$ -axis), it cannot create the necessary torque-resistance in order to dissipate the residual torque.

### B. Valve Turning with Moving Base

In the next set of experiments the base of the manipulator, that is mounted on a wheeled table, is moved manually to create a disturbance during the turning task. These oscillatory disturbances simulate the dynamics of currents in the underwater environment. In this experiment, the duration of the valve turning process is set to  $30 \text{ s}$  to give enough time to the operator to disturb the manipulator. The recorded force and torque data during the valve turning with perturbed base are shown in bottom sub-figures in Figures 7 and 8. The perturbation is generated from  $3 \text{ s}$  up to  $33 \text{ s}$  of the overall time.

## V. LEARNING OF REACTIVE BEHAVIOR

In robotic valve turning in the real world, a sudden movement of the arm can endanger both the valve and the manipulator. Also, if the robot exerts huge and uncontrolled amount of force/torque during the turning phase, it may break

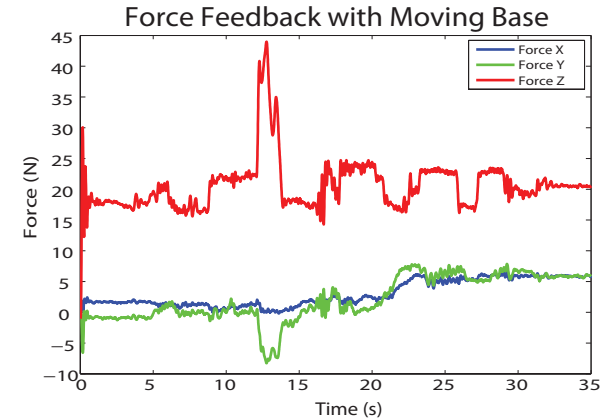
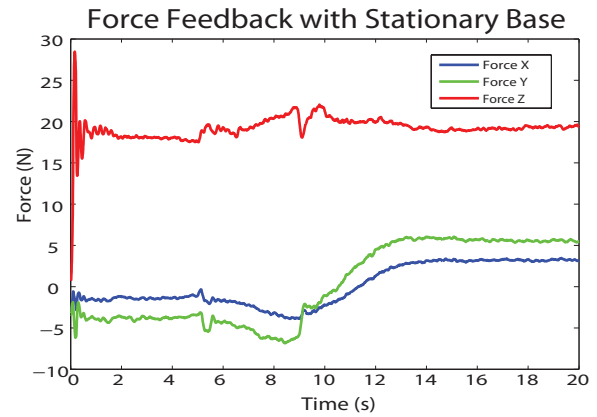


Fig. 7: End-effector forces during the valve turning task.

the valve off. In order to prevent such behaviors and developing a more autonomous and reliable system, a reactive decision maker system is designed. This system, which is a Reactive Fuzzy Decision Maker (RFDM), evaluates the dynamic behavior of the system and regulates the robot's movements reactively.

We chose fuzzy systems because they are based on linguistic rules and the parameters that specify the membership functions have clear physical meanings. Also, there are methods to choose good initial values for the parameters of a fuzzy system [18].

The RFDM system in our previous research [8], monitors the relative movement between the valve and the end-effector and generates decisions according to the defined linguistic rules. One of the drawbacks of the previous reactive system is that, it is independent of the distance between the gripper and the valve. This means that despite the distance between the gripper and the valve, the robot shows identical behaviors. The proposed RFDM in this paper, on the other hand, comprises two more inputs. One is the distance between the gripper and the valve. This extra information gives the RFDM the capability to behave more adaptively. For instance, when the gripper is about to grasp the valve, the new RFDM generates more watchful decisions and increases the sensitivity of the robot's movements with respect to the disturbances. The other input is the force/torque values applied to the gripper and reacts to the uncertainties. For

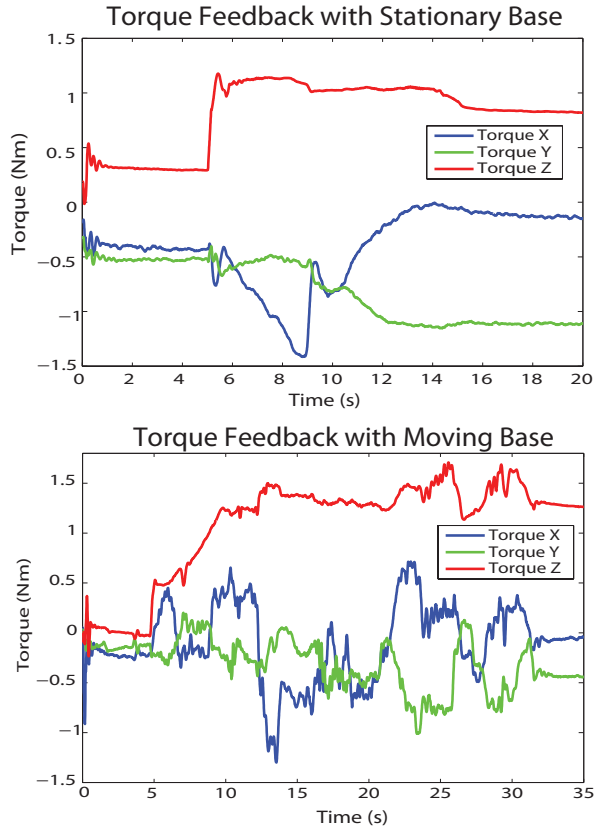


Fig. 8: End-effector torques during the valve turning task.

instance, RFDM retracts the arm when it observes a sudden increase in force/torque during the turning phase.

#### A. Design of the Fuzzy System

The proposed fuzzy system comprises three inputs: *a*) the distance between the gripper and the valve (the norm of the distance vector); and *b*) the relative movement between the valve and the gripper (in  $x-y$  plane); *c*) the forces and torques applied to the valve from the gripper.

All the inputs are first normalized in range  $[0, 1]$  and then are sent to the RFDM system. The third input is provided by the F/T sensor which has a sampling interval equal to  $1\text{ ms}$ . The output of the sensor consists of three force and three torque elements. In this case, the torque is multiplied by a factor to be numerically comparable to the value of the force. The normalizing equation is as follows:

$$\gamma = \frac{\|\mathbf{F}\| + \beta \|\mathbf{T}\|}{F_{max}} \quad (4)$$

where  $\gamma \in [0, 1]$ ,  $\beta = 10$  is a constant factor used to level-off the range of values between the forces and the torques, and  $F_{max} = 30\text{ N}$  is set as the maximum threshold.

Monitoring the relative movement between the valve and the gripper, the system can detect oscillations with different amplitudes and frequencies. For instance, if the end-effector is reaching the valve, and the system senses an oscillation with say *Medium* amplitude the fuzzy system reacts to that by halting the arm. To simulate such behavior in the

experiments, the operator manually moves the table of the robot back and forth. Moreover, considering the distance between the gripper and the valve, the system can change its behavior adaptively. For example, if the gripper is *Far* from the valve, even in the presence of a disturbance, the robot still moves towards the valve. On the other hand, if the gripper is in the vicinity of the valve the robot reacts to smaller oscillations and waits or even retracts the arm. Furthermore, measuring the force/torque magnitudes applied to the gripper, generated by colliding either to the valve or other objects, the system reacts according to the defined rules.

The output of the RFDM system is the reactive decision which is a real number in range  $[-1, 1]$ . The sign of the output specifies the direction of the movement (i.e.,  $+$  for going forward and  $-$  for going backward). For instance,  $-1$  means to retract with 100% speed,  $0$  means to stop, and  $1$  means to approach with 100% speed. Therefore, the RFDM system not only decides the direction of the movement, but also specifies the rate of the movement.

In order to design the fuzzy system, we consider the inputs to be  $\mathbf{u} = [u_1, u_2, u_3]^T$  and the output as  $r$ . Firstly,  $N_i$  ( $i = 1, 2, 3$ ) fuzzy sets,  $A_i^1, A_i^2, \dots, A_i^{N_i}$ , are defined in range  $[0, 1]$ , which are normal, consistent, and complete with Gaussian membership functions  $\mu_{A_1^1}, \mu_{A_2^2}, \dots, \mu_{A_i^{N_i}}$ . Then, we form  $N_{rule} = N_1 \times N_2 \times N_3$  ( $3 \times 4 \times 3 = 36$ ) fuzzy *IF-THEN* rules as follows:

$$\text{IF } u_1 \text{ is } A_1^{i_1} \text{ and } u_2 \text{ is } A_2^{i_2} \text{ and } u_3 \text{ is } A_3^{i_3} \text{ THEN } y \text{ is } B^{i_1 i_2 i_3} \quad (5)$$

Moreover, 7 constant membership function in range  $[-1, 1]$  are set for the output. Finally, the TSK fuzzy system is constructed using product inference engine, singleton fuzzifier, and center average defuzzifier [18]:

$$r = \frac{\sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \sum_{i_3=1}^{N_3} y^{i_1 i_2 i_3} \mu_{A_1^{i_1}}(u_1) \mu_{A_2^{i_2}}(u_2) \mu_{A_3^{i_3}}(u_3)}{\sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \sum_{i_3=1}^{N_3} \mu_{A_1^{i_1}}(u_1) \mu_{A_2^{i_2}}(u_2) \mu_{A_3^{i_3}}(u_3)} \quad (6)$$

Since the fuzzy sets are complete, the fuzzy system is well-defined and its denominator is always non-zero. The designed fuzzy system cannot be illustrated in a single 3D plot because it consists of three inputs and one output. We plotted the fuzzy surface for input variables  $u_2$  and  $u_3$  over a single value of the variable  $u_1$ . So each surface in Figure 9 is related to a fixed value of  $u_1$ . It can be seen from Figure 9 that RFDM shows more sensitive and cautious behaviors as the distance to the valve decreases.

#### B. Tuning the Fuzzy System

In order to tune the parameters of the devised fuzzy system, the subconscious knowledge of a human expert is derived. In this case, the human expert knows what to do but cannot express exactly in words how to do it. In order to extract the subconscious knowledge of the human expert, a tutor simulates the effect of the disturbances (e.g., underwater currents) by moving the wheeled table, while the robot tries to reach and turn the valve. Simultaneously, using a

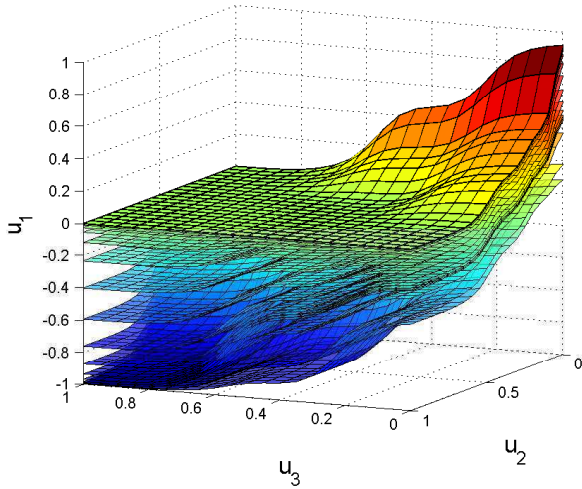


Fig. 9: Fuzzy inference system surface including three inputs ( $u_1, u_2, u_3$ ). The input specifying the distance between the robot and the valve  $u_1$  affects the sensitivity of the designed fuzzy system according to the distance from the valve. Each surface shows a fixed value of the  $u_1$  input for the whole range of the  $u_2$  and  $u_3$  inputs.

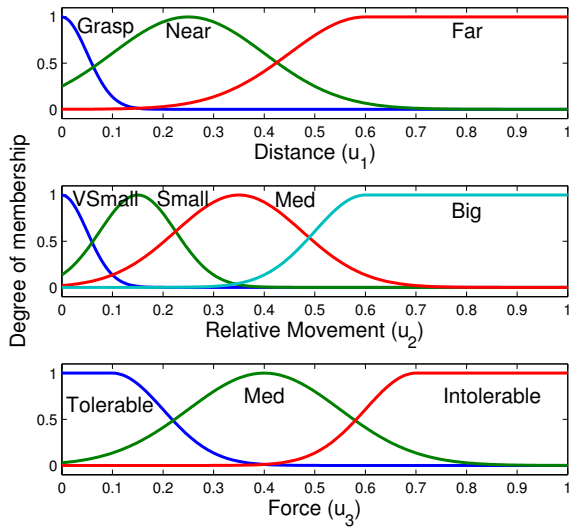


Fig. 10: Fuzzy membership functions defined for each input.

slider button, another tutor regulates the movements of the manipulator while it is following the reproduced trajectory or turning the valve. The tutor applies appropriate continuous commands in range  $[-1, 1]$ , to the system, where  $-1$  means go backward along the trajectory with 100% speed and 1 means go forward along the trajectory with 100% speed. For instance, when the base of the robot is being oscillated say with a *Big* amplitude, the tutor smoothly moves the slider backwards to retract the arm and prevent it from any collision with the valve or the panel. All data, including the position of gripper and the valve, and the tutors commands are recorded during the learning process. The recorded data is then used

to tune the RFDM in off-line mode.

The error between the recorded data from the tutor, which is a fuzzy surface, and the output of the un-tuned fuzzy system which is also a fuzzy surface, is used to make an objective function. The objective function can be minimized using various optimization algorithms. In [8], we applied four different optimization algorithms including gradient-descent, cross entropy method [19], covariance matrix adaptation-evolution strategy (CMA-ES) [20], and modified Price algorithm [21]. The number of optimization parameters for this problem is equal to the number of membership functions multiplied by two (center and standard deviation), plus the number of constant outputs. In our design the number of optimization parameters is equal to 79 (36 Gaussian Membership Functions  $\times$  2 parameters for each Gaussian + 7 constant outputs). In this paper we use CMA-ES for the tuning task, because CMA-ES is typically applied to search space dimensions between three and a hundred [20].

### C. Experimental Result of the Reactive System

In this section, the behavior of the proposed RFDM system is investigated during a real-world valve turning experiment. All three inputs of the reactive system are illustrated in the left side of Figure 11. The first input, the distance, shows that initially the gripper was located far from the valve and gradually approached towards it. The second input, the relative movement, shows that, at some point a relative movement between the gripper and the valve is occurred. The relative movement was created manually by moving the robot's base. And the third input, the force, shows small values during the process but at the end a sudden jump in the force was occurred. The jump in the force magnitude was generated manually by pushing the manipulator towards the valve during the turning phase. The generated decision commands by the RFDM system is plotted in the right subplot of Figure 11. The effect of both the manual oscillation of the base and the manual push on the gripper is observed by RFDM and proper decisions are generated. During the manual oscillation of the base, the RFDM system decreases the rate of the motion towards zero. However, it does not retract the arm because at this point the gripper has not reached the valve yet. Also by sensing the force created by the sudden push, the RFDM system retracts the gripper from the valve.

## VI. DISCUSSION

In our experiments the robot is mounted on a wheeled base which is used to emulate the disturbances. There is a significant level of friction between the wheels and the ground due to the weight of the robot and the base itself. While the robot maintaining the contact with the valve by employing a force controller, such friction prevents the base from moving due to the reactionary forces created between the robot and the valve. In underwater environment in which the robot is floated, the force may exceed the inertia of the vehicle and move it. One solution is to use a combined Vehicle-Arm control strategy in order to overcome

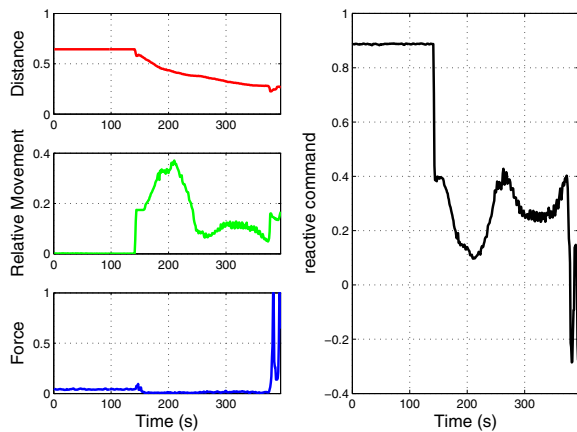


Fig. 11: The recorded set of inputs and the generated decision commands by the RFDM system during a real-world valve turning experiment.

this unwanted behavior. The controller generates more thrust to keep the position of the robot while keeping the contact between the arm and the valve.

Employing the hybrid force/motion controller, the robot's reaction to a stuck valve can happen simultaneously according to three different scenarios. First, due to the desired force specified normal to the valve, the gripper will continue to exert a normal force to maintain contact. Second, due to the desired zero forces and torques along (and around) the  $x$ - and  $y$ -axes, the gripper will automatically adjust its gripping configuration such that the reactionary forces and torques along (and around) these axes are lessened. Lastly, in applying the required torque to turn the valve, the gripper will apply the maximum possible torque around the  $z$ -axis, until the motors get saturated.

## VII. CONCLUSIONS

We have proposed a learning method for reactive robot behavior to deal with the challenging task of autonomous valve turning. The autonomous valve turning consists of two main phases: reaching and turning. Imitation learning is used to learn and reproduce the reaching phase. A hybrid force/motion controller is devised to accomplish the turning phase. In order to increase the autonomy of the system a reactive fuzzy decision maker is developed. This module evaluates the dynamic behavior of the system and modulates the robots movements reactively. The validity and performance of our approach is demonstrated through a real-world valve turning experiment.

## REFERENCES

- [1] D. M. Lane, F. Maurelli, P. Kormushev, M. Carreras, M. Fox, and K. Kyriakopoulos, "Persistent autonomy: the challenges of the PAN-DORA project," *Proceedings of IFAC MCMC*, 2012.
- [2] M. A. Abidi, R. O. Eason, and R. C. Gonzalez, "Autonomous robotic inspection and manipulation using multisensor feedback," *Computer*, vol. 24, no. 4, pp. 17–31, 1991.

- [3] D. A. Anisi, E. Persson, and C. Heyer, "Real-world demonstration of sensor-based robotic automation in oil & gas facilities," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 235–240.
- [4] D. A. Anisi, C. Skourup, and A. Petrochemicals, "A step-wise approach to oil and gas robotics," in *IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, Trondheim, Norway*, vol. 31, 2012.
- [5] M. Orsag, C. Korpela, S. Bogdan, and P. Oh, "Valve turning using a dual-arm aerial manipulator," in *Unmanned Aircraft Systems (ICUAS), 2014 International Conference on*. IEEE, 2014, pp. 836–841.
- [6] N. Alunni, C. Phillips-Graffittin, H. B. Suay, D. Lofaro, D. Berenson, S. Chernova, R. W. Lindeman, and P. Oh, "Toward a user-guided manipulation framework for high-dof robots with limited communication," in *Technologies for Practical Robot Applications (TePRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.
- [7] A. Ajoudani, J. Lee, A. Rocchi, M. Ferrati, E. M. Hoffman, A. Settini, N. G. Tsagarakis, D. G. Caldwell, and A. Bicchi, "Dual arm impedance control with a compliant humanoid: Application to a valve turning task."
- [8] S. R. Ahmadzadeh, P. Kormushev, and D. G. Caldwell, "Autonomous robotic valve turning: A hierarchical learning approach," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 4614–4619.
- [9] A. Carrera, S. Ahmadzadeh, A. Ajoudani, P. Kormushev, M. Carreras, and D. Caldwell, "Towards autonomous robotic valve turning," *Cybernetics and Information Technologies*, vol. 12, no. 3, 2012.
- [10] S. Schaal, A. Ijspeert, and A. Billard, "Computational approaches to motor learning by imitation," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 358, no. 1431, pp. 537–547, 2003.
- [11] M. H. Raibert and J. J. Craig, "Hybrid position/force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, vol. 103, no. 2, pp. 126–133, 1981.
- [12] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *Robotics and Automation, IEEE Journal of*, vol. 3, no. 1, pp. 43–53, 1987.
- [13] T. Yoshikawa and X.-Z. Zheng, "Coordinated dynamic hybrid position/force control for multiple robot manipulators handling one constrained object," *The International Journal of Robotics Research*, vol. 12, no. 3, pp. 219–230, 1993.
- [14] 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.
- [15] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Advanced Robotics*, vol. 25, no. 5, pp. 581–603, 2011.
- [16] S. N. Das and S. K. Das, "Determination of coupled sway, roll, and yaw motions of a floating body in regular waves," *International Journal of Mathematics and Mathematical Sciences*, vol. 2004, no. 41, pp. 2181–2197, 2004.
- [17] R. S. Jamisola, D. N. Oetomo, M. H. Ang, O. Khatib, T. M. Lim, and S. Y. Lim, "Compliant motion using a mobile manipulator: an operational space formulation approach to aircraft canopy polishing," *Advanced Robotics*, vol. 19, no. 5, pp. 613–634, 2005.
- [18] L. Wang, *A Course on Fuzzy Systems*. Prentice-Hall press, USA, 1999.
- [19] R. Y. Rubinstein and D. P. Kroese, *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer, 2004.
- [20] N. Hansen, "The cma evolution strategy: a comparing review," in *Towards a new evolutionary computation*. Springer, 2006, pp. 75–102.
- [21] P. Brachetti, M. D. F. Ciccoli, G. Di Pillo, and S. Lucidi, "A new version of the price's algorithm for global optimization," *Journal of Global Optimization*, vol. 10, no. 2, pp. 165–184, 1997.