Towards Adaptive Robot Motor-Babbling

G. J. Pollayil, M. J. Pollayil, P. Salaris, L. Pallottino and A. Bicchi, *Fellow, IEEE*

***Abstract*—This work takes inspiration from how human infants gradually acquire knowledge about how their body works by performing suitable exploratory actions. With the aim of replicating this important feature also in robotic systems, in this paper we propose a method that combines adaptive robot control, meant to determine both the kinematic parameters and the type of joint (prismatic or revolute), and trajectory optimization, meant to determine the persistent excitation motions required to acquire the maximum amount of knowledge about the robot itself. The ultimate aim is to enable robots to learn and become aware of their physical-mechanical structure. This would favour the application of powerful and predictable model-based control techniques even when only a minimal knowledge of the robot is available (i.e. the number of joints in our case). A preliminary validation is carried out through simulations on simple 1 DoF and 2 DoF robots, showing promising results and potential improvements to be addressed in the future.**

I. Introduction

In the early 50s, the term motor-babbling was introduced in the field of cognitive and developmental psychology [1]. This concept is mainly related to the exploratory actions performed by children during the first months of their life with the aim of learning their own body model and structure by associating the motor commands and the produced actions.

In robotics, a somewhat similar philosophy is adopted by *learning* techniques that allow robots to build models of tasks or of the environment by doing repetitive actions (see [2], [3] and references therein). In [4], the authors developed a motor-babbling-based sensory-motor learning for creating a confidence function of the robot's perception and motor behaviour so as to improve its sensory prediction model.

The incredible potential offered by learning approaches to build accurate models, even for hard-to-engineer tasks, comes mainly at two costs: tedious data harvesting and long training times. Usually, an extensive number of experiments are needed, either from reliable simulations or from the real robot, to collect sufficient data. Then, to achieve a correct fitting of a model to the measured data, time-consuming training sessions are required. This could be quite problematic when dealing with complex physical systems, such as robots, because of the long task execution times and the frequent need for maintenance or repair by human operators [3]. Furthermore, by using generic models, learning techniques often disregard the available basic knowledge about the robot structure (e.g. the kinematic or dynamic model apart from some parameters as well as the number of joints).

On the other hand, the methods of adaptive robot control provide a more computationally efficient and real-time applicable way to control robotic systems even without a perfectly known model but still making use of what little is known about the system. It has been shown in [5] that satisfactory tracking of desired trajectories can be achieved by building an adaptation law for the unknown model parameters, based on the tracking error. In the case of uncertain kinematic or dynamic quantities, adaptive control allows for a precise control of the robot by estimating strictly the necessary robot parameters needed for performing the task, i.e. for zeroing the tracking error. The condition for the correctness of estimation of all the unknown parameters of the robot is the so-called "persistent excitation" properties of the tracked trajectories [6], which were extensively studied by several authors (see, e.g. [7], [8]).

In this work, we provide an initial glimpse towards a new possible approach for learning the model of a robot by motor-babbling using techniques provided by adaptive control. The presented method can be applied to any serial manipulator and requires the only knowledge about the number of joints. Indeed, by building upon the minimum information adaptive control framework presented in [9], we suppose to be completely agnostic of both the geometric parameters and the type of joints of the robot. The kinematic regressor, written using the modified Denavit-Hartenberg (DH) parametrization presented in [9], is employed to perform an optimization-based computation of the persistently exciting trajectories needed for parameter estimation convergence. Thus, tracking such trajectories through adaptive control will allow to progressively learn the true model of the robot from the execution errors.

II. Minimum Information Adaptive Control

To take into account the possibility of minimum or no knowledge about the type of the joints of the robot, [9] proposed a modification of the DH parameters [10] a_i, α_i, d_i and θ_i using an additional parameter

$$r_i = \begin{cases} 0 & \text{if } i^{\text{th}} \text{ joint is prismatic} \\ 1 & \text{if } i^{\text{th}} \text{ joint is revolute.} \end{cases} \quad (1)$$

This work has received funding from the European Union's Horizon 2020 research and innovation program under agreement no. 732737 (ILIAD). The content of this publication is the sole responsibility of the authors. The European Commission or its services cannot be held responsible for any use that may be made of the information it contains.

(Corresponding Author: G. J. Pollayil)

G. J. Pollayil, M. J. Pollayil and A. Bicchi are with the Centro di Ricerca "Enrico Piaggio", Università di Pisa, Pisa 56126, Italy, with the Soft Robotics for Human Cooperation and Rehabilitation, Fondazione Istituto Italiano di Tecnologia, Genova 16163, Italy, and also with the Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Pisa 56126, Italy
(e-mail: georgejose.pollayil@phd.unipi.it; mathewjose.pollayil@phd.unipi.it; antonio.bicchi@unipi.it).

P. Salaris and L. Pallottino are with the Centro di Ricerca "Enrico Piaggio", Università di Pisa, Pisa 56126, Italy, and also with the Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Pisa 56126, Italy
(e-mail: paolo.salaris@unipi.it; lucia.pallottino@unipi.it).

The original D-H parameters are changed as

$$\theta_i = \theta_{0,i} + r_i q_i, \qquad d_i = d_{0,i} + q_i - r_i q_i. \tag{2}$$

Here, q_i is the i^{th} component of the general coordinate vector $q \in \mathbb{R}^{n_q}$, and $\theta_{0,i}$ and $d_{0,i}$ are the i^{th} joint offsets.

As parameters $r_i \in \{0, 1\}$, [9] shows that the differential kinematics of the robot can still be linearly parametrized as[1]

$$\dot{\xi} = J(q, \pi)\dot{q} = Y(q, \dot{q})\pi, \tag{3}$$

where, ξ is the end-effector position, J is the position Jacobian and $\pi = \pi(p)$ is a function of unknown base parameters p. It is noteworthy that, in general, the map $\pi = \pi(p)$ is a set of over-determined equations. For this reason, to find the base parameters p a nearest solution approach, such as least-squares, should be used.

Let $e = \xi_d - \xi$ be the position error with respect to a desired trajectory ξ_d and $\hat{\pi}$ an estimate of the real parameters π. Assuming that the estimated Jacobian $\hat{J} = J(q, \hat{\pi})$ is invertible, by choosing the kinematic control law

$$\dot{q} = \hat{J}^{-1}(\dot{\xi}_d + Ke), \tag{4}$$

and the parameter update law

$$\dot{\hat{\pi}} = -QY^{\mathsf{T}}e, \tag{5}$$

where, K and Q, are symmetric and positive-definite matrices, it can be proved that the error e converges asymptotically to zero. The proof follows the steps of the traditional kinematic adaptive control and can be found in [9].

As it typically happens with adaptive control, the zeroing of the residual error $\tilde{\pi}$ cannot be guaranteed for all q_d. Indeed, $\tilde{\pi}$ will evolve inside a manifold described by the equation $Y(q_d, \dot{q}_d)\tilde{\pi} = 0$. The condition for the convergence of $\tilde{\pi}$ requires that q_d is persistently exciting.

III. Trajectory Optimization

A desired trajectory q_d is persistently exciting [11], if the following integral

$$G_Y(t_f, q_d, \dot{q}_d) = \frac{1}{T}\int_0^{t_f} Y^{\mathsf{T}}(q_d, \dot{q}_d)Y(q_d, \dot{q}_d)\, dt \tag{6}$$

is full rank. The goal of estimating the true parameters could also be seen as a problem of state observability of the system $\dot{\pi} = 0$ from the measured output $y = \dot{\xi} = Y\pi$. This problem is completely observable if the associated observability Gramian, which has the same form as (6), is full rank. Of course, by maximizing some norm of G_Y, (e.g. the smallest eigenvalue), it is possible to determine, among all persistently exciting trajectories, the most persistently exciting one.

The desired trajectory ξ_d should be computed so that the corresponding q_d complies with the full-rank condition of G_Y. Our approach consists in finding the ξ_d that maximizes the smallest eigenvalue λ_{min} of (6). This approach draws inspiration from techniques, such as the *optimal perception-aware trajectory generation* presented in [12].

At the present stage, we choose to optimize over q_d instead of ξ_d in order to avoid performing inverse kinematics during the optimization. We hence parametrize $q_d = q_d(\phi, t)$, where ϕ are parameters that define the geometry and the timing law along the whole trajectory. This could be done by using e.g. Fourier series, B-Spline, Lissajous curves, etc. Once an optimal trajectory in the joint space (solution to the maximization) is found, it shall be transformed to the Cartesian space by using the best available knowledge of the forward kinematics model with the initial estimate of the unknown parameters $\hat{\pi}_0$.

To find the most persistently exciting trajectory we consider the following optimization problem

$$\begin{aligned} \phi^\star = &\arg\max_{\phi} \left(\lambda_{min}\left(G_Y(q_d, \dot{q}_d)\right)\right) \\ &\text{s.t.} \quad l_\phi \le \phi \le u_\phi\,. \end{aligned} \tag{7}$$

The bounds on ϕ are related to the physical limits of the robot. Optimization (7) can be solved by using any state-of-the-art algorithm and allows to get an "optimal" trajectory $q_d(\phi^\star, t)$ (sub-optimal if local maximum is reached) which is persistently exciting: indeed, while this trajectory is being tracked using the adaptive control (4) and the estimation law (5), the robot parameters estimation error $\tilde{\pi}$ will converge to zero.

This notwithstanding, in practice, also the choice of the gains K and Q plays a fundamental role in the convergence of $\tilde{\pi}$ in a reasonable time. This is a classic duality in the problem of adaptive control: there should always be a trade-off between exploration and exploitation.

IV. Simulations

The optimization problem (7) presented in the previous section is solved in MATLAB/Simulink© by using the interior-point method within *fmincon*. For parametrizing the desired trajectory to be tracked by the controller, we choose Fourier series where the i^{th} component of the desired joint trajectory q_d is

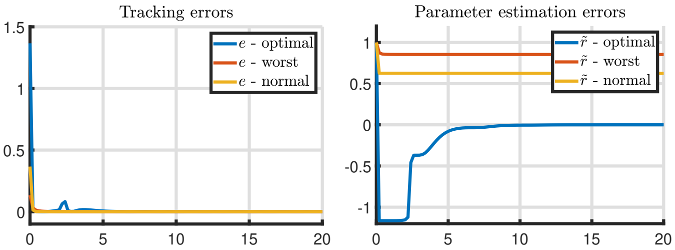
$$q_{i_d}(\phi, t) = \sum_{l=1}^{L}\left[a_{il}\sin(l\omega_f t) - b_{il}\cos(l\omega_f t)\right] + q_{i0}. \tag{8}$$

For the sake of reducing computational complexity, we limit the number of harmonics to one ($L = 1$). Hereinafter, we present two different examples: a single degree of freedom (DoF) and a two DoF robots. In the following, by "optimal" or "worst" trajectory we indicate a trajectory that maximizes or minimizes $\lambda_{min}(G_Y)$, respectively.

A. Single DoF Robot

As a trivial example, consider a simple manipulator with a single joint, also used as "toy example" in [9]. We suppose that the only unknown parameter is the joint type r and we wish to learn it while tracking a trajectory on the horizontal

[1] Indeed, $\cos(r_i q_i) = 1 - r_i - r_i\cos(q_i)$ and $\sin(r_i q_i) = r_i \sin(q_i)$ in the transformation matrices coming from the DH convention.



(a) Tracking errors (b) Parameter estimation errors

Fig. 1: Tracking and parameter estimation errors for the single DoF robot in the cases of a “worst”, ‘normal” and ‘optimal” trajectories.

axis. By making use of (2) and setting $\theta_0 = 0$, the horizontal position of the tip, the related Jacobian and regressor are

$$x(q, r) = (1 - r)(q + d_0) + r d_0 c_q \quad (9)$$
$$J(q, r) = 1 - r(1 + d_0 s_q) \quad (10)$$
$$Y(q, \dot{q}) = [\, \dot{q} \quad -\dot{q}(1 + d_0 s_q) \,]. \quad (11)$$

Note that the parameters vector was chosen as $\pi = [1 \quad r]^\intercal$ and we estimate only the second component. We suppose the real robot to have a revolute joint ($r = 1$) and we pretend to have an erroneous initial knowledge by considering the joint to be prismatic: the estimate $\hat{r}$ of the joint type is hence initialized to 0. The initial condition of the real robot is $q(0) = 5.75\,\mathrm{rad}$ ($q(0) = -0.53 + 2\pi\,\mathrm{rad}$, which gives $x(0) = 0.87\,\mathrm{m}$) and $d_0 = 1\,\mathrm{m}$. Fig. 1 shows the evolution of both the tracking error and the parameter estimation error when different desired trajectories are tracked: the “worst” one in red, the “optimal” one in blue and the one used in [9], i.e. $x_d(t) = 0.5\cos(1\frac{\mathrm{rad}}{\mathrm{s}} \cdot t)\,\mathrm{m}$, hereafter named “normal”. Fig. 2 contains some sequences showing the execution of the “optimal” and the “worst” trajectories by the real robot (in black). In the same Fig. 2 also an “estimated” robot is drawn (in blue) using the estimated parameter $\hat{r}$. The “worst” trajectory happens to be a constant reference $x_d(t) = 1\,\mathrm{m}$; thus, we have a situation of point-to-point control from $x(0)$. In this case, it can be seen that, while the tracking error e converges almost instantaneously to zero, the estimation error of the joint parameter r changes slightly for a very short interval of time and then remains unchanged (see also Fig. 2b where the “estimated” joint remains prismatic). The same can be noticed in Fig. 1 in the case of a “normal” (not “optimal”) desired trajectory On the other hand, while tracking the “optimal” persistently exciting trajectory (blue), both the tracking and the parameters estimation errors quickly converge to zero (see Fig. 2a).

Additionally, for understanding the influence of signal amplitude and frequency on the estimation, we performed also optimizations with some parameters fixed while varying others: in particular, 1) with fixed amplitudes a_{11} and b_{11} and 2) with fixed angular frequency ω_f. In the first case, as expected, an increase in the angular frequency ω_f of the computed trajectory led to better convergence of the parameter estimation error $\tilde{r}$. This is shown in Fig. 3a, where the fixed amplitude values are $a_{11} = 0.4\,\mathrm{m}$ and $b_{11} = 0.4\,\mathrm{m}$ while, for the “optimal” trajectory $\omega_f = 2\,\frac{\mathrm{rad}}{\mathrm{s}}$ and for the “worst” trajectory $\omega_f = 0.1\,\frac{\mathrm{rad}}{\mathrm{s}}$. These values are the bounds that were imposed on the parameter ω_f for the optimization. In the other case, when fixing the angular frequency, an increase in the amplitude of the signals led to better estimation of the parameter. This was also an expected result and it is shown in Fig. 3b where we kept $\omega_f = 1\,\frac{\mathrm{rad}}{\mathrm{s}}$ and the “optimal” trajectory had $a_{11} = 0.3\,\mathrm{m}$ and $b_{11} = 0.3\,\mathrm{m}$ and the “worst” trajectory had amplitudes almost near to zero.

B. Two DoF Robot

In this subsection, we consider a two DoFs robot. As in the previous case, we suppose that the only unknown parameters are the joint types and that we are provided with the other geometric variables (this is mainly done for reducing computational complexity of the simulations). Indeed, the presented approach is able to deal with the identification of all the robot parameters even though complexity increases considerably when using the algorithm in [9] to compute the linear parametrization. The explicit expressions of the tip position, Jacobian and regressor are not reported here for the sake of simplicity and space.

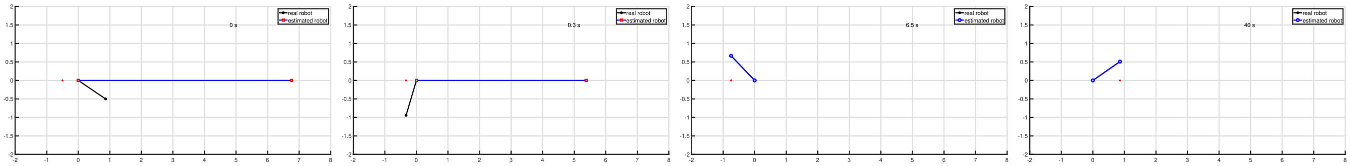
We suppose that the real robot is a two revolute joints planar manipulator, i.e. $r_1 = 1$ and $r_2 = 1$. We initialize the parameters estimates as $\hat{r}_1 = 1$ and $\hat{r}_2 = 0$. In other words, we assume to only know the type of the first joint. The initial conditions of the real robot are $q(0) = [-0.84 \quad 0.05]^\intercal\,\mathrm{rad}$.

By applying our optimization method we found the following “optimal” trajectory, which has the following parameters: $a_{11} = 1.09$; $b_{11} = -0.10$; $a_{21} = 0.72$; $b_{21} = -0.33$; $q_{01} = -0.12$; $q_{01} = -0.29$; $\omega_f = 1.91$. Along the “optimal” trajectory, the estimation of the joint type parameters r_1 and r_2 converges to the real values (see Fig. 4a). Differently from the 1 Dof robot, in this case we used the MATLAB function *lsqnonlin* to estimate the base parameters $p = [r_1\, r_2]^\intercal$ from the map $\pi(p)$. We found also a “worst” trajectory, for which it can be seen (Fig. 4b) that the parameters remain almost unchanged. In accordance with what was seen in the 1 DoF example, also here, the “worst” trajectory consists of a very low signal amplitude and frequency.

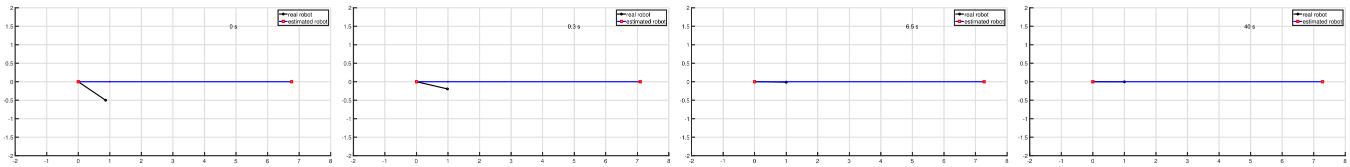
V. Conclusion and Future Works

In this work, we combined adaptive control with trajectory optimization with the aim of proposing a minimum information universal controller that “learns” the model of the robot in order to apply on it strong model-based controls.

To this end, we employed a more generic parametrization of the robot kinematics, obtained from a modification of the Denavit-Hartenberg parameters, which includes also a binary parameter to take into account a possible lack of knowledge about the type of the joint. We then applied kinematic adaptive control with a persistently exciting reference trajectory, which is found by solving an optimization problem using the condition of persistent excitation found in literature. Through simple

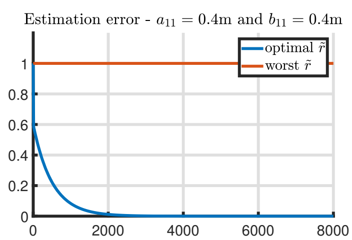


(a) When tracking the “optimal” trajectory, the estimated robot “converges” to the real revolute joint robot.

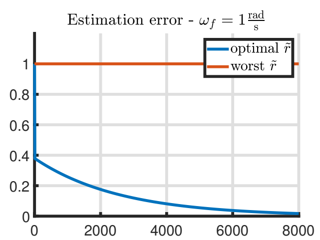


(b) When tracking the “worst” trajectory, the estimated robot remains prismatic even if the true one has a revolute joint.

Fig. 2: Snapshot of the tracking task along “optimal” and “worst” trajectories for the singled DoF robot.

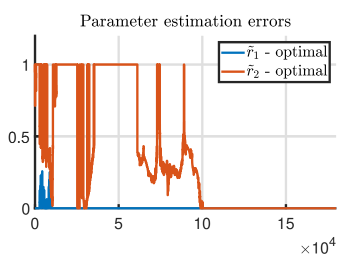


(a) Fixed amplitude

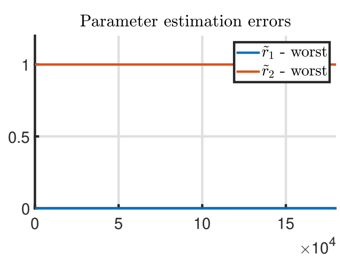


(b) Fixed angular frequency

Fig. 3: Parameter estimation errors for the Single DoF robot in the cases of a ‘worst”, ‘normal” and ‘optimal” trajectories with fixed amplitude or angular frequency of the trajectory.



(a) “Optimal” trajectory



(b) “Worst” trajectory

Fig. 4: Parameter estimation errors for the two DoF robot in the cases of “optimal” and ‘worst” trajectories.

simulations, the proposed framework was found to be valid and to guarantee a satisfactory inference of the robot parameters in reasonable time. However, more simulations and experiments on real robots are necessary to truly assess the efficiency of the approach.

Besides, even though our method has the advantage of being theoretically straightforward, some practical issues might arise because of the computational complexity due to the numeric integration of the convolution integral of the regressor. There are also some other short comings, such as the choice of parametrizing q_d instead of ξ_d. Indeed, future works might include addressing the issue of numerical complexity, also by optimizing directly the end-effector trajectory. We also wish to extend the framework to the case of dynamic adaptive control and to the situation of lack of knowledge even of the number of joints of the robot.

ACKNOWLEDGMENTS

The authors sincerely express their gratitude to Tobia Marcucci and Cosimo Della Santina for the invaluable help and suggestions.

REFERENCES

[1] J. Piaget and M. Cook, *The origins of intelligence in children*. International Universities Press New York, 1952, vol. 8, no. 5.
[2] J. Peters, S. Vijayakumar, and S. Schaal, “Reinforcement learning for humanoid robotics,” in *Proceedings of the third IEEE-RAS international conference on humanoid robots*, 2003, pp. 1–20.
[3] J. Kober, J. A. Bagnell, and J. Peters, “Reinforcement learning in robotics: A survey,” *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238–1274, 2013.
[4] R. Saegusa, G. Metta, G. Sandini, and S. Sakka, “Active motor babbling for sensorimotor learning,” in *2008 IEEE International Conference on Robotics and Biomimetics*. IEEE, 2009, pp. 794–799.
[5] J.-J. E. Slotine and W. Li, “Adaptive robot control: A new perspective,” in *26th IEEE Conference on Decision and Control*, vol. 26. IEEE, 1987, pp. 192–198.
[6] J.-J. E. Slotine and W. Li, “On the adaptive control of robot manipulators,” *The international journal of robotics research*, vol. 6, no. 3, pp. 49–59, 1987.
[7] S. Boyd and S. S. Sastry, “Necessary and sufficient conditions for parameter convergence in adaptive control,” *Automatica*, vol. 22, no. 6, pp. 629–639, 1986.
[8] M. Gautier and W. Khalil, “Exciting trajectories for the identification of base inertial parameters of robots,” *The International journal of robotics research*, vol. 11, no. 4, pp. 362–375, 1992.
[9] T. Marcucci, C. Della Santina, M. Gabiccini, and A. Bicchi, “Towards minimum-information adaptive controllers for robot manipulators,” in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 4209–4214.
[10] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: modelling, planning and control*. Springer Science & Business Media, 2010.
[11] J.-J. E. Slotine and W. Li, “Adaptive manipulator control: A case study,” *IEEE transactions on automatic control*, vol. 33, no. 11, pp. 995–1003, 1988.
[12] P. Salaris, M. Cognetti, R. Spica, and P. R. Giordano, “Online optimal perception-aware trajectory generation,” *IEEE Transactions on Robotics*, vol. 35, no. 6, pp. 1307–1322, 2019.