

Derivative-free online learning of inverse dynamics models

Romeres, D.; Zorzi, M.; Camoriano, R.; Traversaro, S.; Chiuso, A.

TR2018-201 March 15, 2019

Abstract

This paper discusses online algorithms for inverse dynamics modelling in robotics. Several model classes including rigid body dynamics (RBD) models, data-driven models and semi-parametric models (which are a combination of the previous two classes) are placed in a common framework. While model classes used in the literature typically exploit joint velocities and accelerations, which need to be approximated resorting to numerical differentiation schemes, in this paper a new “derivative-free” framework is proposed that does not require this preprocessing step. An extensive experimental study with real data from the right arm of the iCub robot is presented, comparing different model classes and estimation procedures, showing that the proposed “derivative-free” methods outperform existing methodologies.

IEEE Transactions on Control Systems Technology

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Derivative-free online learning of inverse dynamics models

Diego Romeres^{*}, Mattia Zorzi[†], Raffaello Camoriano^{‡,◊}, Silvio Traversaro[◊] and Alessandro Chiuso[†]

Abstract—This paper discusses online algorithms for inverse dynamics modelling in robotics. Several model classes including rigid body dynamics (RBD) models, data-driven models and semiparametric models (which are a combination of the previous two classes) are placed in a common framework. While model classes used in the literature typically exploit joint velocities and accelerations, which need to be approximated resorting to numerical differentiation schemes, in this paper a new “derivative-free” framework is proposed that does not require this preprocessing step. An extensive experimental study with real data from the right arm of the iCub robot is presented, comparing different model classes and estimation procedures, showing that the proposed “derivative-free” methods outperform existing methodologies.

Index Terms—Inverse dynamics, Robotics, Rigid Body Dynamics, Learning, Online methods, Semiparametric models, Derivative-free methods, Gaussian processes, Marginal Likelihood optimisation

I. INTRODUCTION

Robotic platforms, such as industrial and service robots, are becoming more and more popular and are prospected to pervade our everyday life in the near future. A key requirement for such systems is that they should be able to *safely* interact with the environment, possibly including humans, while *performing robustly and efficiently* certain assigned tasks; in order to do so, they have to adapt their behavior to variable conditions.

Among other technological challenges, this requires new mathematical tools for data-driven modeling while accounting for inherent uncertainty and changing conditions (*learning*). While accomplishing user assigned tasks, the robotic system acquires new data which needs to be processed *online* to adapt these models.

In this paper we shall be concerned with *online learning* of so called *inverse dynamics models*: these models have joint trajectories as inputs and joint torques as outputs. They are widely used for model-based control in robotic applications to improve the tracking performances leading to high accuracy and low control gains [1], [2], [3], [4]; see also the survey

[5] for a comprehensive overview. Inverse dynamics models are also very important for detection and estimation of contact forces, see for instance [6].

Typically, inverse dynamics models are obtained from first principles, using the physics of rigid body dynamics (RBD), [7]. This results in a parametric model that depends on some physical parameters (masses, lengths, etc.). Unfortunately, accurate knowledge of these physical parameters may often not be available; for this reason an RBD model should be estimated from data. The main advantage of the parametric approach is that in principle it provides a global relationship between inputs and outputs. However, the parametric model strongly relies on several assumptions and may be rather inaccurate when these assumptions are not perfectly satisfied (frictions, nonlinearities, non perfectly rigid links, etc.), [8], [7].

Alternatively, a nonparametric black-box model can be estimated from experimental data using machine learning techniques such as Gaussian Process regression [9]. The nonparametric framework has the advantage of not requiring unrealistic assumptions, but it comes at the price of being local in nature: the model can only be expected to reflect the systems dynamics in a “neighbourhood” of the trajectories already seen during the learning phase. In the context of system identification for linear dynamical systems a comparison between the capability of parametric and nonparametric approaches has been presented in [10].

To exploit the advantages of both estimation techniques, semiparametric models have been recently introduced as a combination of RBD and nonparametric models, as for instance in [11], [12].

In order to endow robotic systems with the ability to adapt to changing conditions, algorithms should be able to process data online, while taking advantage of knowledge already acquired, in the spirit of so-called transfer learning [13], [14]. Feasibility of online learning strongly hinges on the possibility to keep the computational complexity and memory storage requirements bounded as the number of data grows with time. To this propose several approaches have been proposed in the literature [15], [16], [17], [18], [19], [20], [21] including algorithms explicitly developed to process data online, such as [22], which selects a sparse subset of training data points, and the local Gaussian process regression approach proposed in [2].

In [23] the complexity is kept constant approximating the kernel function using the so-called “random features”, [24], an approximation which will be exploited in this paper following also [22], [25], [26]. It is worth noting that other

This work has been supported by the FIRB project “Learning meets time” (RBF12M3AC) funded by MIUR.

^{*} Mitsubishi Electric Research Labs (MERL), Cambridge, MA, USA (e-mail: romeres@merl.com). Part of the work has been done at [†]

[†] Dept. of Information Engineering, University of Padova, Via Gradenigo 6/b, 35131, Padova, Italy (e-mail: {zorzimatt, chiusoa}@dei.unipd.it)

[‡] LCSL - IIT@MIT, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[◊] iCub Facility, Istituto Italiano di Tecnologia, Via Morego 30, Genoa 16163, Italy (e-mail: {raffaello.camoriano, silvio.traversaro}@iit.it)

approximation techniques are available in the literature, for instance: subset of data approximation, subset of regressors approximation, conditional approximations, Nystrom approximation and relevance vector machine approach, see [27]. Approximation of the kernel is equivalent to choosing a finite (and fixed) basis expansion for the unknown model; this expansion can be exploited to tackle estimation using Kalman filtering techniques [28], [29] and [30], [31]. Online nonparametric learning using Gaussian processes calls also for online tuning of the kernel function, see for instance [32], [33] where marginal likelihood optimisation is used in the context of online system identification.

The main contributions of this paper are as follows:

- Various models (parametric, nonparametric and semi-parametric) proposed in the literature [11], [12], [25] are placed in a common framework through the so-called semiparametric models. This common framework is exploited to compare theoretically the semiparametric model with RBD mean (SP) and the semiparametric with RBD kernel (SPK) and to show (see Section III for details) that the latter is to be preferred, complementing the findings from the experimental evaluation. Online learning is performed, following [25], exploiting the random features approach.
- A new “derivative-free” modelling framework, which avoids the use of numerical derivatives, is proposed. In robot modelling, this framework contrasts the errors introduced into the system by computing numerically the velocities and accelerations starting from the measured (noisy) positions.
- A thorough experimental study, based on real data from the iCub robot, is undertaken; classical methods as well as the newly proposed derivative-free methods are compared and analysed, both in terms of adaptation capabilities (how fast the algorithms can learn the dynamical models after a change in the experiential conditions) as well as in terms of steady state error. The experimental results show that the derivative-free methods proposed in this paper outperform classical schemes based on numerical derivatives. In doing so Cross Validation and Marginal Likelihood optimisation methods are compared for estimating the kernel hyperparameters, suggesting that the latter should be preferred to the former.

The paper is organized as follows. In Section II the problem of inverse dynamics modeling is formalized. In Section III parametric, nonparametric and semiparametric models are introduced, while Section IV deals with kernel approximation which allows for the use of online algorithms. Section V introduces learning methods to avoid the use of numerical derivatives. In Section VI the different online algorithms are tested in the inverse dynamics estimation of the robotic platform iCub. Finally, in Section VII conclusions and future works are drawn.

II. PROBLEM STATEMENT

Assume we are given a robotic system with n degrees of freedom (DoF), and denote with $q(t) \in \mathbb{R}^n$, $t \in \mathbb{R}$, the free

coordinates describing the robot configuration. Starting from the laws of physics, which account for gravity, apparent forces, frictions and so on, it would in principle be possible to write a (direct) dynamical model which, having as inputs the torques $y(t)$ acting on the robot joints, outputs the free coordinates (joint positions) $q(t)$. This is the so called “direct dynamics”.

However, for the purpose of control design, it is of interest to know which torques $y(t)$ should be applied in order for the free coordinates to follow the law $q(t)$. This can be achieved for example using inverse dynamics models, which can be exploited, e.g., to determine the feedforward joint torques $y^d(t)$ which should be applied in order for the free coordinates to follow the law $q^d(t)$; see Figure 1. Clearly, closing a high

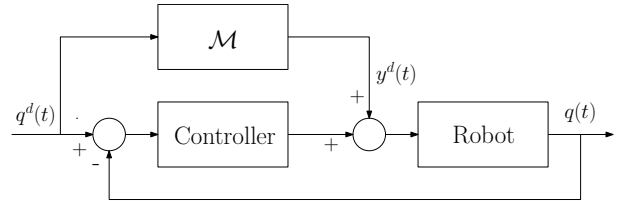


Fig. 1. Schematic for robot motion control.

performance loop quests for an accurate inverse dynamics model \mathcal{M} . While control is an important application of inverse dynamics models, it is certainly not their unique use. For instance, inverse dynamics models find important applications in modelling, detecting and estimating contact forces, see e.g. [6].

In practice an *inverse dynamics model* should be learned from a set of measured data $(y(t), q(t))$, with $t = 1 \dots N$; this should be possibly performed online, updating the model as new data become available. An overview of model structures used for inverse dynamics modeling will be provided in Section III and the approximations involved in online learning discussed in Section IV.

Remark 2.1 (Causality of Inverse Dynamics Models):

Strictly speaking, such models are not proper: the present torques depend on the joint positions, velocities and accelerations which corresponds to having knowledge about future temporal instants. In this paper we shall either assume velocities and accelerations are given, or causal approximations will be considered, where the present torques only depend on present and past joint trajectories. If such models have to be used to build feedforward control laws, it is well possible that the future (desired) trajectories are known and thus one can build, starting from data, non-causal models. On the other hand, when online identification is to be performed (which means that the underlying system is time-varying) there is the delay in updating the time-varying model. However, time-variability should be slow and, as such, it should not be a significant drawback.

Remark 2.2 (Stability of Inverse Dynamics Models): A model of a mechanical system as a map from joint torques $y(t)$ (the inputs) to joint angular velocities $\dot{q}(t)$ (the outputs), is passive. In the linear case this implies that the transfer from input to outputs is positive real and thus it is minimum phase. Nonlinear extensions of this concept are indeed possible,

i.e. a passive nonlinear system has stable zero dynamics, see e.g. [34]. Therefore, when considering inverse dynamics for (passive) mechanical systems, stability is not an issue. Stability of the inverse model may become critical when non minimum-phase systems are considered (e.g. non-collocated systems); it would thus be possible to consider acausal and stable functions of the reference trajectory to build a (stable) inverse dynamics model. This would be compatible with online implementation provided the reference trajectory is known in a finite and sufficiently long future horizon. We shall not consider the latter case in this paper.

The typical assumption that joint positions $q(t)$, joint velocities $\dot{q}(t)$, and joint accelerations $\ddot{q}(t)$ are measured considerably simplifies the identification procedure; the latter are stacked in the so-called “input location” vector

$$x(t) := [q^\top(t) \dot{q}^\top(t) \ddot{q}^\top(t)]^\top \in \mathbb{R}^m \quad (1)$$

with $m = 3n$. Under this assumption, the inverse dynamics of a rigid body can be written in a linear form once a suitable overparametrization is introduced, see Section III-A. Unfortunately, measuring velocities and accelerations is often unrealistic and one has to resort to numerical differentiation schemes, which may be prone to large errors in the presence of measurement noise. In Section V, we shall analyze an alternative model which only exploits the past history of the joint positions. In this way, \mathcal{M} is a model with the past history of the joint positions as input and with the applied torques as output. As we shall see this choice compares favourably with standard approaches in the literature.

III. MODEL CLASSES

In this Section we briefly review the typical model classes used to learn the inverse dynamics of a robot, that is the linear parametric model, whose structure is given by the physics, the nonparametric model, whose structure is learnt from the data, and the semiparametric models which are a combination of the first two models. Throughout this Section, Gaussian processes are indexed in \mathbb{R}^m , take values in \mathbb{R}^n and are zero mean. The symbol $e(t)$ denotes a zero mean white Gaussian noise with covariance matrix $\sigma^2 I_n$. Most of the models introduced in this Section are equipped with a so-called *hyperparameters vector*. We address the problem of estimating this vector in Section IV.

A. Linear parametric model

The rigid body dynamics (RBD) of a robot is described by the equation

$$y(t) = M(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + G(q(t)) \quad (2)$$

where $M(q(t))$ is the inertia matrix of the robot, $C(q(t), \dot{q}(t))$ the Coriolis and centripetal forces and $G(q(t))$ the gravity forces, [7], [35]. The terms on the right hand side of (2) can be rewritten as $\psi^\top(x(t))\pi$ which is linear in the robot (base) inertial parameters vector $\pi \in \mathbb{R}^p$. The map $\psi : \mathbb{R}^m \rightarrow \mathbb{R}^{p \times n}$ is the known RBD regressor which is a nonlinear function of

the input locations vector described by $x(t)$. Therefore, the RBD model is equivalent to

$$y(t) = \psi^\top(x(t))\pi + e(t) \quad (3)$$

where $e(t)$ includes the nonlinearities of the robot that are not modeled by the rigid body dynamics (e.g. actuator dynamics, frictions, etc.).

Since the RBD model is physics-based, it should, in principle, describe the robot dynamics for all the desired trajectories, leading to good generalization and global approximation properties. A known issue of this model (see e.g., [8]) is that the problem of determining π from measured data $y(t)$ is usually ill posed and the matrix $\psi(x(t))$ is rank deficient. Possible solutions in system identification are either the design of efficient experiments to collect data sufficiently rich to excite the highest number of modes of the system or dedicated experiments which are good to estimate parameters separately.

Another drawback of such model class is that it is prone to undermodeling (as it may not capture non-linear friction effects, effects of non-rigidities etc.) that may ultimately result in severely biased estimated models.

This model will be denoted as “P” (Parametric).

B. Nonparametric Model

Following the “Gaussian Process” framework [9], the robot inverse dynamics can be modeled postulating that:

$$y(t) = f(x(t)) + e(t) \quad (4)$$

where $f(\cdot)$ is a Gaussian process with covariance function (i.e. kernel function)¹ $\text{cov}[f(x(t)), f(x(s))] = K(x(t), x(s))$:

$$K(x(t), x(s)) = \rho^2 K_G(x(t), x(s)) I_n. \quad (5)$$

Several choices are possible for the kernel K (see e.g. [9], [36]), which can be exploited to encode specific model structures. However, in this paper we shall only consider K to be the Gaussian kernel²

$$K_G(x(t), x(s)) = e^{-\frac{\|x(t) - x(s)\|^2}{2\tau}} \quad (6)$$

where the hyperparameter τ is the kernel width. The latter is a typical choice to correlate the input locations for learning the inverse dynamics, [23], [11], [12]. ρ^2 plays the role of scaling factor. The vector $\eta := [\rho^2 \tau \sigma^2]$ is referred to as *hyperparameters vector*. This model will be denoted as “NP” (NonParametric).

This class of models is known to have high flexibility and prediction performance (see e.g. [9]) since the dynamics are extrapolated directly from the experimental data, without making any unrealistic approximation on the physical system (e.g. assuming linear frictions models, ignoring the dynamics

¹To be precise, the kernel function and the covariance function coincide up to the scaling factor. However, to ease the exposition we will follow this abuse of terminology.

²In order to avoid scaling issues, each component of the input location vector $x(t)$ is normalised to have unit standard deviation. Some preliminary tests showed worse performance when using an independent width for each component of $x(t)$, which is probably to be attributed to overfitting due to the large number of hyperparameters.

of the hydraulic actuators, etc.). Nevertheless, nonparametric models deteriorate their performance when predicting unseen data which are far (in the Euclidean metric) from those visited in the training dataset.

C. Semiparametric model with RBD mean

The semiparametric model is an attempt to take advantage of the global property of the parametric model as well as of the flexibility of the nonparametric model; the first possibility is to embed the former as a mean term into the latter. Thus, the inverse dynamics is modeled as

$$y(t) = \psi^\top(x(t))\pi + f(x(t)) + e(t) \quad (7)$$

where π is the vector of inertial parameters, $\psi(x(t))$ is the RBD regressor and $f(\cdot)$ a Gaussian process with kernel function as in (5).

At this point two alternatives are possible. The first and most principled one is to treat π as an unknown hyperparameter. Model (7) with this hypothesis will be denoted as ‘‘SP’’ (SemiParametric). In this case $\eta := [\pi \rho^2 \tau \sigma^2]$.

A suboptimal alternative but often applied in the literature is to assume that π is known, here denoted by $\hat{\pi}$. This could be possibly estimated using some preliminary experiment as in [11] or estimated via Least Squares as in [25] or it can be given from some expert knowledge. Model (7) with this hypothesis will be denoted as ‘‘SP2’’. In this case $\eta := [\rho^2 \tau \sigma^2]$.

D. Semiparametric model with RBD kernel

An alternative possibility for combining the parametric and nonparametric models is to incorporate the RBD structure in the kernel, [11]. The debate as to whether one should account for prior knowledge as a mean term or as a structured kernel is often encountered in Gaussian Process regression. We shall come back to the relation between these two alternatives in Proposition 3.1 and the discussion which follows that proposition.

Therefore, the inverse dynamics is modeled as

$$y(t) = f(x(t)) + e(t) \quad (8)$$

where $f(\cdot)$ is a Gaussian process with kernel function

$$K(x(t), x(s)) = \gamma^2 \psi(x(t))^\top \psi(x(s)) + \rho^2 K_G(x(t), x(s)) I_n. \quad (9)$$

This model will be denoted as ‘‘SPK’’ (SemiParametric Kernel). The hyperparameters vector is $\eta := [\gamma^2 \rho^2 \tau \sigma^2]$.

It is worth noting that in the case $\rho = 0$ we obtain the parametric model (3) where the inertia parameters vector is now modeled as a Gaussian random vector with zero mean and covariance matrix $\gamma^2 I_p$.

Remark 3.1: Note that equation (9) is derived assuming π is a zero mean Gaussian vector with variance $\gamma^2 I$. Of course if prior knowledge was available on the expected scale of the components of π , this could be included in the prior variance introducing a diagonal scaling matrix D , so that $\pi \sim \mathcal{N}(0, \gamma^2 D)$.

We have also tested an alternative version in which the diagonal matrix $\gamma^2 D$ has been estimated via Marginal Likelihood together with all other hyperparameters. However the overall performance is worse than when assuming $\pi \sim \mathcal{N}(0, \gamma I)$ and estimating only γ , which is probably due to the fact that this extra freedom results in a less favourable bias-variance tradeoff.

The semiparametric model with RBD kernel (SPK) is connected with the RBD mean (SP) model introduced in Section III-C. To make this connection sharp we shall refer to both the minimum variance estimators (see Section IV) obtained with these models, as well as to the log-likelihood functions which can be used to estimate the hyperparameters (see Section IV-B). In particular, let us stack the available data $y(t)$, $t = 1 \dots N$ in the vector \mathbf{y} and stack correspondingly the regressors $\psi^\top(x(t))$ and $f(x(t))$ in the matrix Ψ and vector \mathbf{f} , respectively. Moreover, we define $\mathbf{K}(\mathbf{x}, \mathbf{x}) = \text{cov}[\mathbf{f}, \mathbf{f}]$, $L_{SPK}(\mathbf{y})$ and $L_{SP}(\mathbf{y})$ as the negative marginal log-likelihoods $-2 \log p_\eta(\mathbf{y})$ of data \mathbf{y} as a function of hyperparameters η for models SPK and SP, respectively, and

$$\hat{L}_{SP}(\mathbf{y}) := \min_{\pi} L_{SP}(\mathbf{y}) = L_{SP}(\mathbf{y})|_{\pi=\hat{\pi}_{WLS}} \quad (10)$$

as the *profile* marginal log-likelihood, where $\hat{\pi}_{WLS}$ denotes a suitable weighted least squares estimate of π (see Appendix A for the precise definition of $\hat{\pi}_{WLS}$). The following proposition establishes the link between the semiparametric with RBD mean (SP) and semiparametric with RBD kernel (SPK) estimators as the prior on π becomes uninformative (i.e. as $\gamma^2 \rightarrow \infty$), showing that also in this asymptotic regime the two models induce a different marginal likelihood (see eq. (11)). The take-home message of this proposition is that the SPK model better handles uncertainties and should thus be preferred; a more in depth discussion will be provided in Section III-E. The reader is also referred to [37] where a similar discussion has been provided in a completely different context.

Proposition 3.1: Consider the semiparametric models SP in (7) and SPK in (8), their respective negative marginal log-likelihoods $L_{SP}(\mathbf{y})$ and $L_{SPK}(\mathbf{y})$ and the profile marginal likelihood $\hat{L}_{SP}(\mathbf{y})$ defined in (10). Assume that ρ^2, τ, σ^2 are fixed. Then, the following two statements hold:

- (i) the minimum variance estimator of model SP and model SPK coincide when $\gamma^2 \rightarrow \infty$;
- (ii) in the limiting case of $\gamma^2 \rightarrow \infty$ the two log-likelihoods $L_{SPK}(\mathbf{y})$ and $\hat{L}_{SP}(\mathbf{y})$ differ in a nontrivial term $\log(\det(\Psi^\top R^{-1} \Psi))$ which leads to different location in their minima. That is:

$$\lim_{\gamma^2 \rightarrow \infty} L_{SPK}(\mathbf{y}) - \hat{L}_{SP}(\mathbf{y}) - p \log \gamma^2 = \log(\det(\Psi^\top R^{-1} \Psi)) \quad (11)$$

where $R = \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma^2 I$.

Proof: See Appendix A. ■

E. Discussion of Proposition 3.1

Proposition 3.1 shows that, when estimating hyperparameters using the marginal likelihoods (as discussed in IV-B), the two models (SP and SPK) are not equivalent even under

the assumption that a non informative prior on π is used, i.e. $\gamma^2 \rightarrow \infty$. In fact, the two marginal likelihood functions differ by a nontrivial term $\log(\det(\Psi^\top R^{-1} \Psi))$ which may change the location of the minima. In particular the latter term accounts for the uncertainty in estimating the term π .

When the SP model is used, the hyperparameters σ^2, ρ^2, τ are estimated minimising

$$\hat{L}_{SP}(\mathbf{y}) = \log(\det(2\pi R)) + (\mathbf{y} - \Psi \hat{\pi}_{WLS})^\top R^{-1} (\mathbf{y} - \Psi \hat{\pi}_{WLS}).$$

Doing so, σ^2, ρ^2, τ are chosen so as to fit with R the ‘‘sample’’ covariance $(\mathbf{y} - \Psi \hat{\pi}_{WLS})(\mathbf{y} - \Psi \hat{\pi}_{WLS})^\top$. The drawback of this choice is that the component of the variance along Ψ may be underestimated (i.e. R too small along Ψ). On the contrary, when optimising $L_{SPK}(\mathbf{y})$ the term

$$\log(\det(\Psi^\top R^{-1} \Psi)) = -\log(\det(\text{Var}\{\hat{\pi}_{WLS}\}))$$

favour values of σ^2, ρ^2, τ which make $\log(\det(\text{Var}\{\hat{\pi}_{WLS}\}))$ large or, equivalently, R^{-1} small (i.e. R large) along the direction of Ψ .

Clearly this goes in the opposite direction and avoids the risk alluded at above (i.e. R too small along Ψ). Thus, to summarise, the term $\log(\det(\Psi^\top R^{-1} \Psi))$ serves as a correction which accounts for the uncertainty of $\hat{\pi}_{WLS}$. Thus it is fair to say that model SPK deals more ‘‘cautiously’’ with uncertainty than model SP. This is reflected in the simulation results (see e.g. Figure 4) where SPK slightly outperforms SP (both endowed with ML hyperparameters estimation, i.e. SP-ML and SPK-ML).

IV. KERNEL APPROXIMATION AND ONLINE LEARNING

In this Section we review the online learning problem using the model classes of Section III. Since it is well known how to perform online learning using the parametric model (3), we will focus on the other models. On the other hand, model (3) can be understood as a special case, for instance, of model (7). Online learning using nonparametric and semiparametric models can be performed by applying the Gaussian regression framework, [9]. In order to do so, two issues have to be faced which will be addressed in Section IV-A and Section IV-B, respectively. First, following [25], [26], an approximation of the kernel is introduced to allow online learning with constant complexity. Second, two commonly used approaches are very briefly reviewed to estimate the hyperparameters vector η , namely Cross Validation and Marginal Likelihood optimisation.

A. Kernel approximation

A typical route followed in machine learning is to approximate the kernel with a low rank factorisation of the form

$$K_G(x(t), x(s)) \approx \phi^\top(x(t)) \phi(x(s)) \quad (12)$$

where $\phi : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$ is a suitable vector of ‘‘basis functions’’ that have to be properly defined. Ideally, $\phi(x(t))$ should be built from the eigenfunctions of the kernel matrix; however this optimal approximation is computationally unfeasible in a recursive algorithm³. Therefore, we rely on the random

³An exception is the case when the eigenfunctions of the kernel are available analytically.

features approximation proposed in [24] and exploited for online inverse dynamics modelling in [23], [25], [26].

According to the random feature approximation, the basis functions vector for the Gaussian kernel is defined as⁴

$$\phi(x) = \frac{1}{\sqrt{d}} \begin{bmatrix} \cos\left(\frac{\omega_1^\top x}{\tau}\right) & \dots & \cos\left(\frac{\omega_d^\top x}{\tau}\right) \\ \sin\left(\frac{\omega_1^\top x}{\tau}\right) & \dots & \sin\left(\frac{\omega_d^\top x}{\tau}\right) \end{bmatrix}^\top \quad (13)$$

with $w_k \sim p(\omega) = \exp(-\|\omega\|^2/2)/(\sqrt{2\pi})^m$, $k = 1, \dots, d$. Notice that if $d \rightarrow \infty$ then the approximation is almost surely exact. As a consequence, the parameter d has to be chosen to trade-off accuracy of the approximation and computational complexity. Finally, it is worth noting that $\phi(x)$ depends on the width of the Gaussian kernel τ .

Using the approximation in (12) for model (8) is equivalent to consider the approximation

$$f(x(t)) \approx \psi^\top(x(t)) \pi + [\phi^\top(x(t)) \otimes I_n] \pi_{NP} \quad (14)$$

where π and π_{NP} are independent Gaussian random vectors with zero mean and covariance matrices $\gamma^2 I_p$ and $\rho^2 I_{2d}$, respectively. Finally, by defining

$$\begin{aligned} \varphi^\top(x(t)) &= [\psi^\top(x(t)) \quad \phi^\top(x(t)) \otimes I_n] \\ \theta &= [\pi^\top \quad \pi_{NP}^\top]^\top \end{aligned}$$

the approximation of model (8) is

$$y(t) = \varphi^\top(x(t)) \theta + e(t). \quad (15)$$

It is possible to show that all the models in Section III can be approximated with a model in the form (15). A derivation of all these approximations can be found in [26].

B. Online learning

Next, we address the problem of online learning. For simplicity of exposition, we consider the NP model in (4). It is well known that, given data $(y(t), x(t))$, $t = 1 \dots N$, the minimum variance estimator of f can also be expressed as the solution of the Tikhonov regularization problem, [9],

$$\hat{f} = \underset{f \in \mathcal{H}}{\operatorname{argmin}} \frac{1}{\sigma^2} \sum_{t=1}^N \|y(t) - f(x(t))\|^2 + \frac{1}{\rho^2} \|f\|_{\mathcal{H}}^2 \quad (16)$$

where f belongs to the reproducing kernel Hilbert space \mathcal{H} , [39], with reproducing kernel function K_G and norm $\|\cdot\|_{\mathcal{H}}$.

It is not difficult to see that the Tikhonov regularization problem (16) takes a similar form for all the models of Section III. The solution of (16) takes the general form

$$\hat{f}(x) = \sum_{t=1}^N \alpha_t K(x, x(t)) \quad (17)$$

⁴This construction is based on the fact that any bounded and positive semidefinite kernel is the Fourier transform of a non-negative and integrable function (Bochner’s theorem), which thus induces a probability distribution (the function $p(\omega)$ below equation (13)). See [24] for details. Another interpretation is that $p(\omega)$ represents (up to a constant factor) the power spectral density [38] of the stationary kernel K_G .

where α_t 's are uniquely determined from $(x(1), y(1)), \dots, (x(N), y(N))$. Accordingly, the solution of (16) is unique regardless of the excitation properties of $x(t)$. Moreover, it is possible to prove that α_t 's take finite value, so that \hat{f} maps bounded trajectories into bounded torques/forces. It is worth noting that the number of coefficients α_t coincides with the number of the data points/samples, i.e. N , making the estimator intractable for an online (recursive) solution. On the other hand, using the approximated model (15), the minimum variance estimator of θ is the solution to the following Regularized Least Squares problem

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \frac{1}{\sigma^2} \sum_{t=1}^N \|y(t) - \varphi^\top(x(t))\theta\|^2 + \theta^\top W \theta \quad (18)$$

where W is the inverse kernel matrix induced by the approximation. Its optimal solution can be computed recursively, whenever new data becomes available, through the well known Recursive Least Squares algorithm, see e.g [40, Chapter 11] and [41]. In practice, the implementation of this algorithm uses Cholesky-based updates [42], which have robust numerical properties. The computational complexity of each update is $O(\bar{p}^2)$ where \bar{p} is the dimension of vector θ ; the cost of evaluating the model output is $O(\bar{p})$ for each output and thus $O(\bar{p}n)$ if n wrenches (torques/forces) are to be computed. This computational complexity is compatible with online implementation for state-of-the-art computation facilities.

C. Hyperparameter vector estimation

Nonparametric and semiparametric models are characterized by the hyperparameters vector η , which is not known and needs to be estimated from the data. Typically, two approaches are considered to address this problem.

A first possible method is called *Cross Validation* (CV). The goal is to obtain an estimate of the prediction capability of the model on future data for different choices of the hyperparameters vector η . Hold-out cross validation approach is a possible version of CV, see e.g. [43, Chapter 6]. The dataset is split in two, the training set and the validation set. The optimal η is given by optimizing an estimate of the mean squared error in the validation set

$$\hat{\eta}_{CV} = \underset{\eta \in \Omega}{\operatorname{argmin}} \widehat{\operatorname{MSE}}(\eta). \quad (19)$$

In practice this approach is limited to the estimation of a small number of hyperparameters since the minimization in (19) is typically performed by gridding the search space Ω .

A second method is offered directly by the Gaussian regression framework. The *marginal likelihood* (ML), denoted by $p_\eta(\mathbf{y})$, expresses the likelihood of the hyperparameters vector η on the data $\mathbf{y} := [y(1)^\top \dots y(N)^\top]^\top$, once the parameter θ has been integrated out. Under model (15), the latter can be computed in closed form, as discussed in [9]. To conclude, the optimal η is then given by solving

$$\hat{\eta}_{ML} = \underset{\eta \in \Omega}{\operatorname{argmax}} p_\eta(\mathbf{y}). \quad (20)$$

In Section VI we shall thoroughly compare on real data these two approaches for hyperparameter estimation. The experimental results, in this case, show that Marginal Likelihood optimisation outperforms Cross Validation.

V. DERIVATIVE-FREE LEARNING

In Section III-A we have seen that the rigid body dynamics suggests that the inverse dynamics is described by a map from the input locations vector $x(t) = [q^\top(t) \dot{q}^\top(t) \ddot{q}^\top(t)]^\top$ composed by joint positions-velocities-accelerations to the joint torques applied to the n joints. This hypothesis has been exploited in all the model classes in Section III. However, it is often the case that joint velocities and accelerations cannot be measured from the robot; rather they are estimated using numerical differentiation schemes from the (noisy) measurements of the joint positions. As a consequence, failure to properly handle noise in the measurement may severely hamper the final solution. This is a very well known and highly discussed problem, see e.g., [7], [8], [44], [1], [5] and it is usually partially addressed by ad-hoc filter design. However, this requires users' knowledge and experience in tuning the filters' parameters.

In this Section we propose a new methodology, which avoids the use of numerical pre-differentiation issues, by learning the inverse dynamics without using the estimated velocities and accelerations. We shall thus assume that only joint torques $y(t)$ and joint positions $q(t)$, $t = 1 \dots N$, are measured. Let

$$q(t^-) := [q_1^\top(t^-) q_2^\top(t^-) \dots q_n^\top(t^-)]^\top \in \mathbb{R}^{(M+1)n}, \quad (21)$$

and

$$q_i(t^-) := [q_i^\top(t) q_i^\top(t-1) \dots q_i^\top(t-M)]^\top \in \mathbb{R}^{M+1}, \quad (22)$$

be the vector of the past joint positions and the vector of the past of the i -th joint position, respectively, in the time window $[t-M, t]$ where M , sufficiently large, has been fixed. Our aim is to model the inverse dynamics as a map from the past joint positions $q(t^-)$ to the joint torques. In particular, we shall postulate that output $y(t)$ can be written as a non-linear function of a "features vector" $\xi(t) := [\xi_1^\top(t) \xi_2^\top(t) \dots \xi_n^\top(t)]^\top$, defined as a linear function of past measurements $q_i(t^-)$

$$\xi_i(t) = R q_i(t^-) \quad (23)$$

where $R \in \mathbb{R}^{k \times (M+1)}$. Using the features (23) can be seen as a reduced rank regression problem. We shall discuss later on the choice of the number k (rows of R), i.e. the number of features which are used by model (24) below⁵. The choice of k could also be performed resorting to Bayesian arguments, as for instance done in [45] in a similar low-rank regression problem in the context of sparse-low rank dynamic network models.

In particular, we shall model the inverse dynamics with the NP model

$$y(t) = f(\xi(t)) + e(t) \quad (24)$$

⁵For example, in the standard case, where we have the input locations vector, the features are $k = 3$: joint positions, velocities and accelerations.

where $f(\cdot)$ is a Gaussian process with kernel function

$$K(\xi(s), \xi(t)) = \rho^2 K_G(\xi(s), \xi(t)) I_n. \quad (25)$$

Then, the online learning algorithm is similar to the one sketched in Section IV. The only difference is that the “standard” input locations vector $x(t)$ in (1) is replaced with $\xi(t)$ in (23). Special cases which one may consider are $R = I$ so that $\xi(t)$ in (23) coincides with the past measurements (i.e. no dimensionally reduction is performed); as an alternative (we shall come back to this later on) the rows of the matrix R may compute, using causal filtering operations, an approximation of the derivatives of $q(t)$ (i.e. $\dot{q}(t)$, $\ddot{q}(t)$ and possibly higher order derivatives). We may thus say that the approach considered in this section generalises what seen in the previous ones.

In the following, the matrix R in (23) will be described by a set of hyperparameters that have to be estimated from the data; for instance, for the NP model we shall have $\eta := [\rho^2 \tau \sigma^2 \delta^\top]$, where δ is the vector containing the hyperparameters of R . Models P, SP, SP2 and SPK can be derived in a similar way.

We shall now consider three possible choices (but of course many others would be possible) for the structure of the matrix R .

A. Derivative-free features

The simplest choice is to take $R = I_{M+1}$ that is the features vector coincides with the M past measured joint positions. Thus $k = M + 1$. These features will be denoted as “Derivative-Free” (DF). As an alternative it is possible to choose

$$R = \begin{bmatrix} r_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & r_{M+1} \end{bmatrix} \quad (26)$$

with $r_i \in \mathbb{R}, i = 1, \dots, M + 1$. That is, the features vector is a weighted version of the past measured joint positions. These features will be denoted as “Derivative-Free Weighted” (DFW).

B. Derivative-free features with reduced rank

An alternative choice is to take the number of features k smaller than $M + 1$. For instance, the physics suggests that the right number of features should be equal to 3 (position, velocity, acceleration). Therefore, the role of the features is to compress the useful information available in $q(t^-)$ so as to render the learning procedure more robust. This means that

$$R = \begin{bmatrix} r_1^\top \\ r_2^\top \\ \vdots \\ r_k^\top \end{bmatrix}, \quad \|r_i\| = 1 \quad (27)$$

where $k \ll M$, $r_i \in \mathbb{R}^{M+1}$ and the vector of hyperparameters is $\delta := [r_1, \dots, r_k]$, i.e. R is fully parameterised. In practice it makes also sense to impose some constraints in the

r_i 's. In particular $r_1^\top := [1 \ 0 \ \dots \ 0]$ so that the first feature is the measured position. It would also be reasonable to enforce other constraints, for instance imposing that R is an orthogonal projection. Notwithstanding its importance, we shall not delve further into this issue.

The number of features k is chosen by the user and it will be subject of empirical analysis in Section VI. These features will be denoted as “Derivative-Free Reduced rank” (DFR).

Notice, that the features DFR include all the possible linear and causal numerical differentiation and filtering operations. The price of this generality is that a large number of hyperparameters has to be estimated, i.e. $(k - 1)(M + 1)$. As it is well known in optimization, this might lead to local minima problems. In order to overcome this issue one might resort to regularization techniques on the hyperparameters or to set appropriate initial conditions.

C. Structured derivative-free input locations with reduced rank

The last idea is to consider only $k = 3$ features: the first will be the position $q(t)$ while the other two will attempt to estimate explicitly velocities and accelerations. We know that the joint velocities and accelerations can be computed by a first order backward difference and by a second order backward difference, respectively, both filtered by a first order low pass filter, that is

$$\begin{aligned} \dot{q}_i(t) &\approx \frac{1 - z^{-1}}{T_s} \frac{1}{1 - \beta_1 z^{-1}} q_i(t) \\ \ddot{q}_i(t) &\approx \frac{1 - 2z^{-1} + z^{-2}}{T_s^2} \frac{1}{1 - \beta_2 z^{-1}} q_i(t) \end{aligned}$$

where z^{-1} is the backward shift operator, $T_s > 0$ is the sampling time and $0 < \beta_1, \beta_2 < 1$ represent the poles of the filters.

We resort to a partial fraction decomposition to rewrite the above expressions as a function of $q(s^-)$, that is:

$$\dot{q}_i(t) \approx \alpha_1 q_i(t) + \sum_{t=1}^M \alpha_1 \beta_1^{t-1} (\beta_1 - 1) q_i(s - t) \quad (28)$$

$$\begin{aligned} \ddot{q}_i(t) &\approx \alpha_2 q_i(t) + \alpha_2 (\beta_2 - 2) q_i(s - 1) \\ &\quad + \sum_{t=2}^M \alpha_2 \beta_2^{t-2} (\beta_2^2 - 2\beta_2 + 1) q_i(s - t) \end{aligned} \quad (29)$$

where $\alpha_1 = 1/T_s$ and $\alpha_2 = 1/T_s^2$. Here, we exploited the fact that a (stable) low-pass filter can be approximated by finite impulse response (FIR) filter with length M , where the latter is chosen sufficiently large. Accordingly, we have

$$R = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ \alpha_1 & \alpha_1 (\beta_1 - 1) & \dots & \alpha_1 \beta_1^{t-1} (\beta_1 - 1) & \dots \\ \alpha_2 & \alpha_2 (\beta_2 - 2) & \dots & \alpha_2 \beta_2^{t-2} (\beta_2^2 - 2\beta_2 + 1) & \dots \end{bmatrix}. \quad (30)$$

These features are denoted as “Derivative-Free Structured Reduced rank” (DFSR) since the features have a structure to resemble the joint positions, velocities and accelerations.

A nice property of this characterization is that the number of hyperparameters in R is small, in this case $\delta = [\alpha_1, \alpha_2, \beta_1, \beta_2] \in \mathbb{R}^4$, independently of the length of the past temporal lags M , which can be arbitrarily chosen.

VI. INVERSE DYNAMICS LEARNING ON iCUB

In this section, we shall provide a thorough experimental evaluation of all the models discussed in Section III, also endowed with derivative-free input locations as discussed in Section V, for online learning of the inverse dynamics; all the algorithms have been tested using real data from the iCub robot [46], which is a full-body humanoid robot with 53 degrees of freedom.

In this work we consider inverse dynamics modelling of the right arm. Therefore the free coordinates $q(s) \in \mathbb{R}^4$ are the angular positions of the 3 shoulder joints and of the elbow joint, for a total of 4 degrees of freedom. When needed, joint positions have been numerically differentiated to obtain joint velocities and accelerations by the Authors of [25] using the standard `adaptWinPolyEstimator`⁶ module of the open source iCub project. The differentiation procedure consists in applying a time-varying linear filter based on the work [47], that is actually implemented at a higher rate before downsampling the signals.

The outputs $y(s) \in \mathbb{R}^6$, are the 3 forces and 3 torques components measured by the six-axes force/torque (F/T) sensors embedded in the shoulder of the iCub’s arm, see Figure 2.

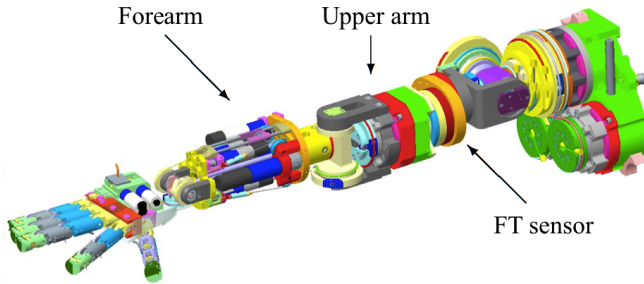


Fig. 2. iCub’s right arm.

Notice that the measured forces/torques are not the applied joint forces and torques and, as such, the model we learn is not, strictly speaking, the inverse dynamics model. Yet, as explained in [48], the feedforward joint torques can be determined from components (forces and torques) of $y(s)$. Indeed, such model has been used in the literature as a benchmark for the inverse dynamics learning, [23], [25].

We consider the two datasets used in [25], corresponding to different trajectories of the end-effector. In the first one (*XY-dataset*) the end-effector tracks circles in the XY plane of radius 10cm at an approximate speed of 6m/s; in the second one (*XZ-dataset*), the end-effector tracks similar circles but in the XZ plane (the Z axis corresponds to the vertical direction, parallel to the gravity force). The two circles are tracked using

the Cartesian controller proposed in [49]. Each dataset contains approximately 8 minutes of data collected at a sampling rate of 20Hz, for a total of 10000 points per dataset. One single circle is completed by the robot in about 1.25 seconds, which corresponds to 25 points.

We shall consider the models described in Section III endowed with standard input locations or derivative-free features and hyperparameters estimated via either the marginal likelihood approach or Cross Validation⁷, the latter has been discussed in [25]. For ease of exposition the models will be denoted with a combination of the following shorthands:

- $P, NP, SP, SP2, SPK$: to indicate the model class, as discussed in Section III, approximated according to Section IV-A.
- ML, CV : to indicate the method used to estimate the hyperparameters, according to Section IV-B.
- $DF, DFW, DFR, DFSR$: to indicate the different derivative-free features, according to Section V. If nothing is indicated, the standard input locations are considered.

For example, we shall denote “NP-ML”, the nonparametric model (4) with hyperparameters estimated through maximization of the marginal likelihood; instead “NP-ML-DFR” shall denote the nonparametric model (24) with hyperparameters estimated through maximization of the marginal likelihood and derivative-free features defined in (27).

The estimation routine has been implemented using Matlab. The RBD regressor ψ for the iCub’s right arm has been computed using the library `iDynTree`, [50]. The Marginal Likelihood has been optimized using the Matlab `fmincon.m` function. The recursive least squares algorithms have been implemented using the `GURLS` library, [51]. The results of all CV methods are obtained using code which has been kindly provided by the authors of [25].

For each model as above, the following online learning scenario is considered (with reference to the general model structure (15)):

- *Initialization*: The first 1000 points in XY-dataset are used to estimate the hyperparameters with one of the two techniques considered, say $\hat{\eta}_{ML}$ and $\hat{\eta}_{CV}$, as well as to compute an initial estimate of the parameter θ , say θ_{1000} .
- *Online Estimation - Stage 1*: The remaining 9000 points of the XY-dataset are used to update online the parameter θ using the recursive least squares algorithm, thus obtaining $\hat{\theta}_t, t = 1001, \dots, 10000$. Let $\hat{\theta}_F^1 = \hat{\theta}_{10000}$ be the final value obtained by this procedure on the XY-dataset.
- *Online Estimation - Stage 2*: The XZ-dataset is split in 5 sequential subsets (*XZ-dataset-i, i = 1, ..., 5*) of 2000 points each (approximately 100 seconds). These subsets are used to evaluate the performance of the online estimators. For each subset, the estimator of θ is always initialized with $\hat{\theta}_F^1$ and updated recursively on the 2000 data of each dataset *XZ-dataset-i, i = 1, ..., 5*.

⁷As discussed in Section IV-B, using the Cross Validation methods is unfeasible when the number of hyperparameters is large; therefore we have not applied validation to the semiparametric model with RBD mean when the mean is to be considered as an hyperparameter nor to the semiparametric model with RBD kernel which has the extra parameter γ^2 .

⁶Available at wiki.icub.org/brain/adaptWinPolyEstimator_8cpp_source.html

Note that the estimators in Stage 2 are initialised using the final estimate from Stage 1, which corresponds to a different motion (XY-dataset). Therefore, the evaluation of the performance in Stage 2 allows us to verify how well the estimators generalise on new unseen data (initial part of the new dataset) as well as how well they are able to learn adapting to a new experimental condition (transient and steady state).

In order to measure the quality of the estimated models, we evaluate the online prediction capability of the estimated models using the following index:

$$\varepsilon_t^{(k)} = \frac{\sum_{s=1}^T (y^{(k)}(t+s) - \hat{y}^{(k)}(t+s|t))^2}{\sum_{s=1}^T (y^{(k)}(t+s))^2} \quad (31)$$

where $\hat{y}^{(k)}(t+s|t)$ is the estimate of the k -th output $y^{(k)}(t+s)$ at time $t+s$ using the model $\hat{\mathcal{M}}_t$ estimated with data up to time t . Note that the test data $y^{(k)}(t+s)$, $s = 1 \dots T$, have not been used to estimate the model $\hat{\mathcal{M}}_t$. Thus, $\varepsilon^{(k)}(t)$ is an average relative error on a future horizon of length T for the output component k . In addition, $\varepsilon^F(t)$ and $\varepsilon^T(t)$ will be the average values of $\varepsilon^{(k)}(t)$ for the 3 forces ($k = 1, 2, 3$) and the 3 torques ($k = 4, 5, 6$), respectively. In all our simulations, T will be equal to 25 (which roughly corresponds to one revolution of the end effector in the XZ plane).

The index (31) can be motivated, for instance, by the possible use of the model in the framework of model predictive control [52], where the final performance of the controller hinges on the predictive capabilities of the model on a given future horizon (say T).

The experimental results will be first presented for models that use numerical derivatives and then for the derivative-free models; a comparison will be eventually provided between the two. For each experiment, we show $\varepsilon^F(t)$ and $\varepsilon^T(t)$ averaged over the 5 subsets XZ-Dataset- i , $i = 1, \dots, 5$.

A. Experimental results using numerical derivatives

In Figure 3 the behavior of $\varepsilon^F(t)$ and $\varepsilon^T(t)$ is presented.

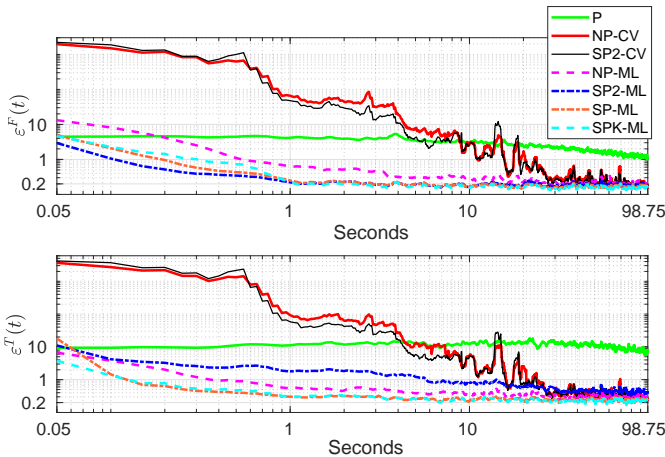


Fig. 3. Average (over the 5 subsets of 100 seconds each) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to an horizon of 1.25 seconds. A log-log scale is used for ease of readability.

The parametric model, P, exhibits a poor performance because

it describes only crude idealizations of the actual dynamics. The algorithms based on Cross Validation (CV) perform significantly worse in the first 60 seconds than those based on Marginal Likelihood (ML) optimisation; this is not unexpected as discussed in [53].

As expected, the nonparametric model, NP-ML, has worse generalization performance (the error is larger in the first few steps) but better adaptation capabilities with respect to model P. The models with the best performance are SP-ML and SPK-ML because they combine the benefit of the parametric approach, i.e. generalization capabilities (good estimation performance at the beginning of the new dataset) and of the nonparametric approach, i.e. learning capabilities (good transient and steady state performance). The estimator SP2-ML should partially inherit these benefits from the SP structure, yet it shows an overall slightly worse performance. This is due to the fact that the first (least squares) step, i.e. the estimation of the linear model, is subject to a strong bias deriving from the unmodeled dynamics. Instead, a sound approach is followed by SP-ML and SPK-ML in which the estimation of the hyperparameters is performed jointly, avoiding such bias. In the steady state these semiparametric models outperform the others; yet the semiparametric SPK-ML performs best both in terms of average as well as distribution, as shown in Figure 4 which reports the boxplots of ε_t^F and ε_t^T in “steady state”, i.e.

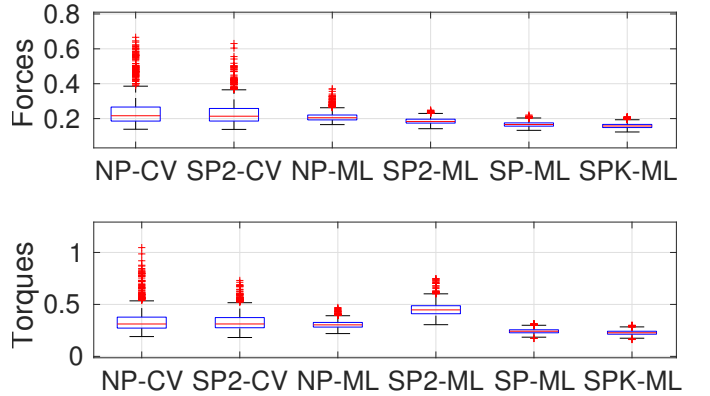


Fig. 4. Boxplots of the steady state (i.e. after 30 seconds, see also Figure 3) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to an horizon of 1.25 seconds.

after the first 30 seconds which is considered to be transient, see Figure 3. The reader is also referred to the discussion in Appendix III-E for a theoretical justification of this latter fact.

B. Experimental results with derivative-free features

The focus of this section is the analysis of the derivative-free models, including the comparison with the models that use numerical derivatives presented in Section VI-A.

The parameters M (number of past temporal lags used to form the features vector $\xi(t)$, see equations (26), (27) and (30), and the number of features, k , in the reduced derivative-free models are set as follows:

- M is fixed equal to 10; larger values have been tested with no significant differences in performance.

- k is fixed to 3. A discussion on this choice as well as some results with different choices can be found later in this Section, see e.g. Figure 7.

The hyperparameters are estimated using marginal likelihood (ML) maximization, since performing Cross Validation via gridding in high dimension is computationally unfeasible.

In the first comparison, the nonparametric methods (NP-ML) is compared to its derivative-free versions NP-ML-DF, NP-ML-DFW, NP-ML-DFR and NP-ML-DFSR.

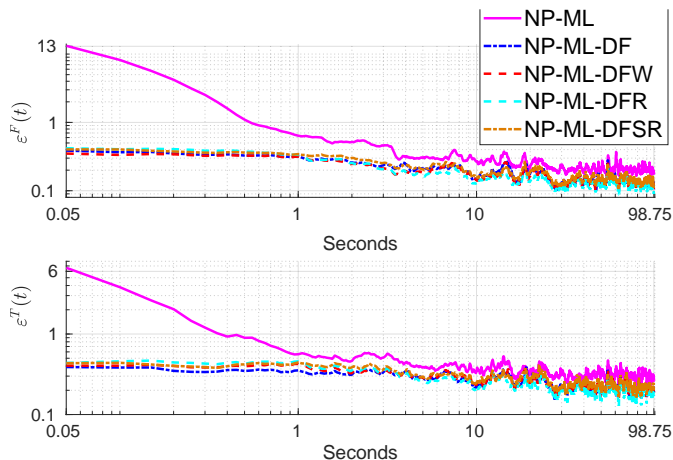


Fig. 5. Average (over the 5 subsets of 100 seconds each) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to a horizon of 1.25 seconds. A log-log scale is used for ease of readability.

In Figure 5 the averaged (over 5 realisations) time evolutions of $\varepsilon^F(t)$ and $\varepsilon^T(t)$ are illustrated. All the nonparametric derivative-free models perform comparably and outperform NP-ML, both in transient (more significant) as well as in steady state. The distribution of the steady state errors is shown using boxplots in Figure 6. It is clear that all the derivative-free methods outperform NP-ML, but also that NP-ML-DFR (which uses a reduced rank derivative-free feature) is the best performing method. This confirms that the dimensionality reduction in equation (23) captures the relevant information and allows to reduce the variance of the estimators.

As anticipated, the results in Figure 5 and Figure 6 are obtained setting $k = 3$ in NP-ML-DSR. The reasons for this choice is that, when using numerical differentiation, the input location vector $x(t)$ contains exactly 3 components (position, velocity and acceleration) for each DoF. It is natural to ask what happens as k changes. In Figure 7 the steady state behaviour⁸ of $\varepsilon^F(t)$ and of $\varepsilon^T(t)$ is analysed as a function of $k = 1, 2, 3, 4$.

It is apparent from Figure 7 that $k = 1$ and $k = 2$ are not sufficient and lead to a considerable performance degradation; in addition, no improvement is obtained increasing k beyond 3 (compare $k = 3$ and $k = 4$ in Figure 7). Larger values (e.g. $k = 5$) have also been tested leading to similar results and have been therefore omitted. The results in Figure 6 show that the performance of NP-ML-DF (which uses $R = I$) is

⁸The transient behaviour has been omitted because it does not add much information with respect to the steady state statistics.

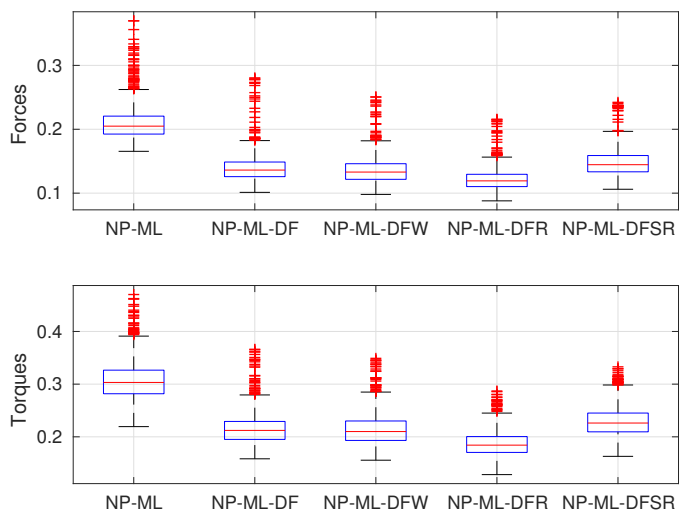


Fig. 6. Boxplots of the steady state (i.e. after 30 seconds, see also Figure 5) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to a horizon of 1.25 seconds.

worse than NP-ML-DFR with $k = 3$. Overall this suggests that, indeed, $k = 3$ is the optimal choice for this specific application. NP-ML-DFR with $k = 3$ performs better than NP-ML-DFSR. Considering the bias-variance trade-off dilemma, the latter method introduces more bias and less variance than the former, suggesting that, in these experimental results, the bias introduced by NP-ML-DFSR is preponderant with respect to the reduction of the variance.

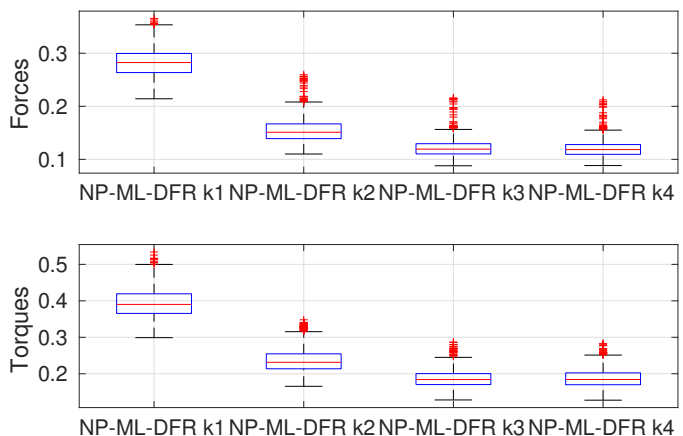


Fig. 7. Boxplots of the steady state (i.e. after 30 seconds) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to a horizon of 1.25 seconds.

Finally, a comparison between the estimators using numerical derivatives and the derivative-free methods (and in particular DFR) is provided for the semiparametric models.

In Figure 8 and Figure 9 the behaviour of $\varepsilon^F(t)$ and of $\varepsilon^T(t)$ for the models SPK-ML-DFR and SPK-ML is illustrated.

The two semiparametric methods have similar initial performance, but SPK-ML has a better learning rate in the transient. This behaviour might be attributed to the parametric component of the model since in SPK-ML-DFR the physical meaning of the features is lost. However, in steady state the semiparametric derivative-free model outperforms the standard

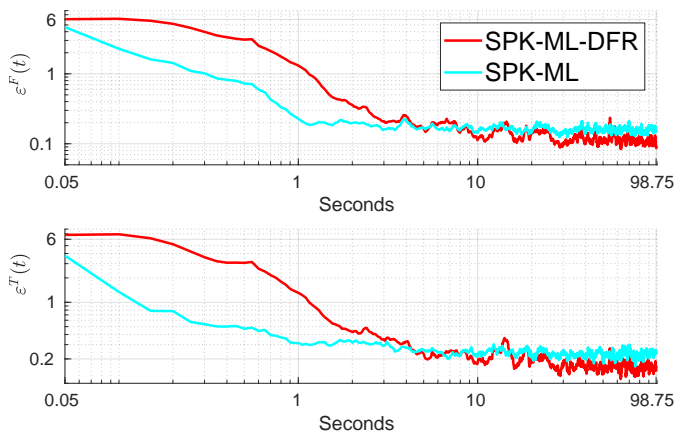


Fig. 8. Average (over the 5 subsets of 100 seconds each) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to a horizon of 1.25 seconds. A log-log scale is used for ease of readability.

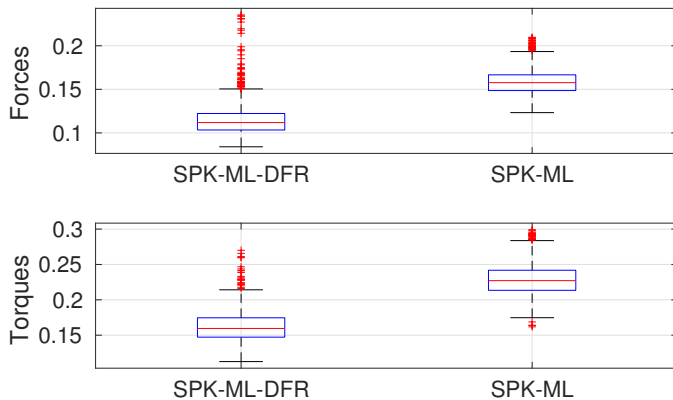


Fig. 9. Boxplots of the steady state (i.e. after 30 seconds, see also Figure 8) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to a horizon of 1.25 seconds.

model with an improvement of about 30% in terms of relative error, see Figure 9.

The results obtained in this section give an empirical evidence that learning the features directly from the past history of the measured joint trajectories is a rather promising direction.

C. Comparison among DFR-like models

Finally a comparison among the models with DFR features, namely, P-ML-DFR, NP-ML-DFR, SP2-ML-DFR and SPK-ML-DFR, is presented in Figures 10 and 11.

The transient performance can be analyzed by observing the first 30 seconds in Figure 10. Similarly to P-ML, also the parametric model P-ML-DFR has very poor performance (note that the relative error is always larger than 1); the poor performance of the parametric models seems to negatively affect also the transient behaviour of the semiparametric models. As a matter of fact, the transient performances of SP2-ML-DFR and SPK-ML-DFR are very poor. The nonparametric model, NP-ML-DFR, definitely outperforms all the other models in terms of transient, suggesting that the parametric models are probably inadequate to capture the dynamic behaviour.

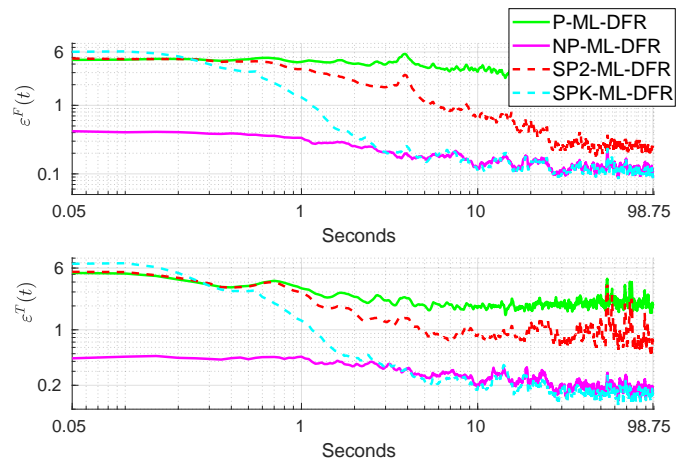


Fig. 10. Average (over the 5 subsets of 100 seconds each) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to a horizon of 1.25 seconds. A log-log scale is used for ease of readability.

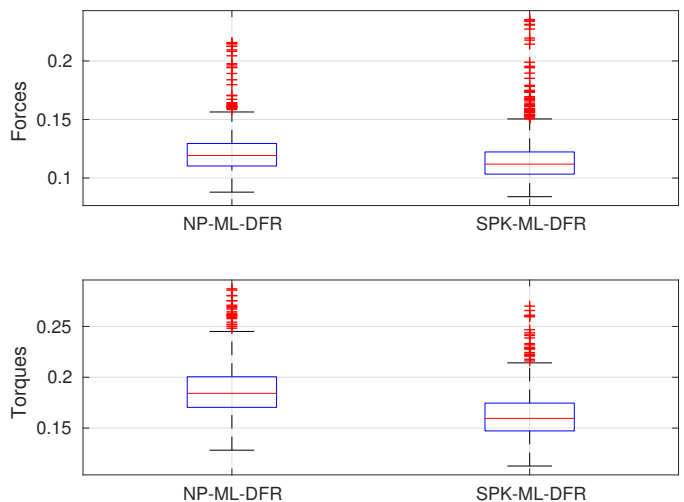


Fig. 11. Boxplots of the steady state (i.e. after 30 seconds, see also Figure 10) of the relative squared prediction errors $\varepsilon^F(t)$ and $\varepsilon^T(t)$, computed with $T = 25$ corresponding to a horizon of 1.25 seconds.

The last 60 seconds of the experiment reported in Figure 10 and the boxplots in Figure 11 (only for NP-ML-DFR and SPK-ML-DFR) illustrate the behaviour at steady state. The suboptimal model SP2-ML-DFR is largely unsatisfactory, which can be probably attributed to the bias error introduced by the parametric component estimated by least squares.

D. Experimental results discussion

We can summarise as follows the findings of our extensive experimental study:

- The derivative-free methods outperform the schemes based on numerical differentiation; this is even more remarkable if one recalls that the numerical derivatives have been computed starting from data at a higher sampling rate, i.e. they have practically used a richer dataset (which ideally should have lead to better noise-rejection properties). The main reason is probably to be attributed to the fact that the derivatives on which the “classical”

models rely are computed using numerical differentiation schemes from the measured positions, which are subject to noise. However, the models (4), (7) and (8) do not account for this noise and operate as if the measurements were correct. Instead, the derivative free method performs a reduction, which can be assimilated to computing derivatives and accelerations, but this reduction is computed as part of the modelling exercise, and as such its effect on wrench prediction is directly accounted for.

- The transient performances of semiparametric models strongly depend on the availability of an accurate (physical) model. Indeed, when using derivative-free methods wherein the extracted features are not guaranteed to approximate velocities and accelerations, the parametric model loses physical meaning. This is confirmed by the poor transient performance of P-ML-DFR, SP2-ML-DFR and SPK-ML-DFR (see Figure 10).
- Depending on whether transient or steady state performance is to be favoured, NP-ML-DFR or SPK-ML-DFR should be preferred respectively (see Figures 10, 11).

VII. CONCLUSIONS

In this paper, several models and algorithms for online inverse dynamics learning have been discussed and compared in a common framework. Parametric and nonparametric methods have been considered, as well as semiparametric models obtained by combining the two. Different strategies, Cross Validation (CV) and Marginal Likelihood optimisation (ML), for estimating the hyperparameters have been compared and also new model structures which do not rely on pre-computed numerical derivatives have been introduced.

All these algorithms have been thoroughly tested on real data from the right arm of the iCub robot; the comparison has been performed both in terms of transient behaviour (how fast algorithms can adapt to a new experimental setup) as well as steady state behaviour.

Overall our experimental validation suggests that:

- The non-parametric model (NP) and the Semiparametric model with RBD mean (SP2) perform better when trained using Marginal Likelihood optimisation rather than Cross Validation algorithms available in [25];
- Semiparametric methods which exploit physical insight, when using numerical derivatives, outperform purely non-parametric structures;
- Derivative-free methods are definitely advantageous w.r.t. ad-hoc methods which rely on numerical differentiation, when applied to purely non-parametric model structures (see e.g. NP-ML-DFR);
- Endowing semiparametric methods with derivative-free schemes is not entirely trivial; these new features do not yield improvements in the parametric model as much as they do in the nonparametric one. As a result the transient behaviour of SPK-ML-DFR is significantly worse than its nonparametric counterpart; instead, SPK-ML-DFR outperforms its nonparametric counterpart in steady state.

This last item calls for further research efforts in the future, which would hopefully allow to fully exploit the benefits of

derivative-free methods coupled with semiparametric model structures. Our future agenda includes further comparison with parametric methods which account for physical consistency of the parameters (see for instance [54]) as well as the implementation of control strategies which exploit the estimated models.

VIII. ACKNOWLEDGMENTS

The authors gratefully acknowledge the other teams involved in the FIRB project ‘‘Learning meets time’’ (RBF12M3AC), in particular iCub Facility group (Francesco Nori, Giorgio Metta) and Lorenzo Rosasco (LCSL-IIT@MIT) for making their data and code available to us.

APPENDIX

A. Proof of Proposition 3.1

First, note that both the SP model and the SPK model can be rewritten as

$$y(t) = \psi^\top(x(t))\pi + f(x(t)) + e(t) \quad (32)$$

where :

- 1) π is an unknown but fixed quantity for the SP model;
- 2) π is a zero mean Gaussian random vector with covariance matrix $\gamma^2 I_p$ for the SPK model.

In both models, $f(\cdot)$ is a Gaussian process with kernel function

$$K(x(t), x(s)) = \rho^2 K_G(x(t), x(s)).$$

A well known connection between Bayes and Fisher (i.e. assuming the parameter π is an unknown but fixed quantity) estimators, is that the latter can be obtained as a limiting case of the former when:

- the parameter π is modeled as a zero mean Gaussian vector with covariance matrix $\gamma^2 I_p$
- the variance of the prior distribution of π is let go to infinity by letting $\gamma^2 \rightarrow \infty$.

This proves the first part of the Proposition. We proceed to prove the second part. Let us stack the available data $y(t)$, $t = 1 \dots N$ in the vector \mathbf{y} and stack correspondingly the regressors $\psi^\top(x(t))$ and $f(x(t))$ in the matrix Ψ and vector \mathbf{f} , respectively, so that models (32) can be written as

$$\mathbf{y} = \Psi\pi + \mathbf{f} + \mathbf{e} \quad (33)$$

where \mathbf{e} is defined with the same rule as \mathbf{y} . Moreover, we define $\mathbf{K}(\mathbf{x}, \mathbf{x}) = \text{cov}[\mathbf{f}, \mathbf{f}]$ and $\mathbf{K}(\cdot, \mathbf{x}) = \text{cov}[f(\cdot), \mathbf{f}]$. The minimum variance estimators of π and f under 2) and given ρ^2, σ^2, τ , are:

$$\begin{aligned} \hat{\pi} &= \text{cov}[\pi, \mathbf{y}] \text{Var}^{-1}[\mathbf{y}] \mathbf{y} \\ &= \gamma^2 \Psi^\top (\gamma^2 \Psi \Psi^\top + \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma^2 I)^{-1} \mathbf{y} \quad (34) \\ \hat{f}(\cdot) &= \text{cov}[f(\cdot), \mathbf{y}] \text{Var}^{-1}[\mathbf{y}] \mathbf{y} \\ &= \mathbf{K}(\cdot, \mathbf{x}) (\gamma^2 \Psi \Psi^\top + \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma^2 I)^{-1} \mathbf{y}. \end{aligned}$$

Defining $R := \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma^2 I$ and using the matrix inversion lemma we have:

$$\begin{aligned} &(\gamma^2 \Psi \Psi^\top + \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma^2 I)^{-1} \\ &= (\gamma^2 \Psi \Psi^\top + R)^{-1} \\ &= R^{-1} - R^{-1} \Psi (\Psi^\top R^{-1} \Psi + \gamma^{-2} I)^{-1} \Psi^\top R^{-1} \end{aligned}$$

so that, from (34)

$$\begin{aligned}\hat{\pi} &= \gamma^2 \left(I - \Psi^\top R^{-1} \Psi (\Psi^\top R^{-1} \Psi + \gamma^{-2} I)^{-1} \right) \Psi^\top R^{-1} \mathbf{y} \\ &= (\Psi^\top R^{-1} \Psi + \gamma^{-2} I)^{-1} \Psi^\top R^{-1} \mathbf{y},\end{aligned}$$

and similarly

$$\begin{aligned}\hat{f}(\cdot) &= \mathbf{K}(\cdot, \mathbf{x}) R^{-1} \left(I - \Psi (\Psi^\top R^{-1} \Psi + \gamma^{-2} I)^{-1} \Psi^\top R^{-1} \right) \mathbf{y} \\ &= \mathbf{K}(\cdot, \mathbf{x}) R^{-1} [\mathbf{y} - \Psi \hat{\pi}].\end{aligned}$$

Clearly, as $\gamma^2 \rightarrow \infty$, we have that $\hat{\pi}$ converges to the *weighted* least squares estimate

$$\hat{\pi}_{WLS} = (\Psi^\top R^{-1} \Psi)^{-1} \Psi^\top R^{-1} \mathbf{y} \quad (35)$$

and \hat{f} converges to

$$\hat{f}(\cdot) = \mathbf{K}(\cdot, \mathbf{x}) R^{-1} [\mathbf{y} - \Psi \hat{\pi}_{WLS}].$$

On the other hand the marginal likelihood function for model (33), under 1), i.e. when π is considered as an unknown parameter, has the form:

$$\begin{aligned}L_{SP}(\mathbf{y}) &= -2 \log(p_\eta(\mathbf{y})) \\ &= \log(\det(2\pi R)) + (\mathbf{y} - \Psi \pi)^\top R^{-1} (\mathbf{y} - \Psi \pi).\end{aligned}$$

When ρ^2, σ^2, τ are kept fixed, the minimization with respect to π can be performed in closed form, and yields exactly the weighted least squares solution (35). However, even for $\gamma^2 \rightarrow \infty$, the marginal likelihoods of \mathbf{y} given the hyperparameters ρ, σ^2, τ under 1) and 2) are different. In fact, if 1) is postulated, and π is solved as above, one obtains the *profile* marginal log-likelihood $\hat{L}_{SP}(\mathbf{y}) := L_{SP}(\mathbf{y})|_{\pi=\hat{\pi}_{WLS}}$

$$\hat{L}_{SP}(\mathbf{y}) = \log(\det(2\pi R)) + (\mathbf{y} - \Psi \hat{\pi}_{WLS})^\top R^{-1} (\mathbf{y} - \Psi \hat{\pi}_{WLS})$$

where the hyperparameters ρ^2, σ^2, τ are hidden in the definition of $R = \mathbf{K}(\mathbf{x}, \mathbf{x}) + \sigma^2 I$.

Instead, if 2) is postulated, the marginal log-likelihood takes the form

$$L_{SPK}(\mathbf{y}) = \log(\det(2\pi \Sigma(\gamma))) + \mathbf{y}^\top \Sigma(\gamma)^{-1} \mathbf{y}.$$

where $\Sigma(\gamma) := \gamma^2 \Psi \Psi^\top + R$. Using, as above, the matrix inversion Lemma on $\Sigma(\gamma)$, the *Sylvester* determinant identity and defining $\hat{\pi} := (\Psi^\top R^{-1} \Psi + \gamma^{-2} I)^{-1} \Psi^\top R^{-1} \mathbf{y}$, we obtain

$$\begin{aligned}L_{SPK}(\mathbf{y}) &= \log(\det(2\pi R)) + \log(\det(I_p + \gamma^2 \Psi^\top R^{-1} \Psi)) \\ &\quad + (\mathbf{y} - \Psi \hat{\pi})^\top R^{-1} (\mathbf{y} - \Psi \hat{\pi}) \\ &\quad + \hat{\pi}^\top \Psi^\top R^{-1} (\mathbf{y} - \Psi \hat{\pi}).\end{aligned}$$

As $\gamma^2 \rightarrow \infty$ we have that $\hat{\pi} \rightarrow \hat{\pi}_{WLS}$ and $\hat{\pi}^\top \Psi^\top R^{-1} (\mathbf{y} - \Psi \hat{\pi}) \rightarrow 0$ so that⁹

$$\begin{aligned}L_{SPK}(\mathbf{y}) &\approx \log(\det(2\pi R)) + \log(\det(I_p + \gamma^2 \Psi^\top R^{-1} \Psi)) \\ &\quad + (\mathbf{y} - \Psi \hat{\pi}_{WLS})^\top R^{-1} (\mathbf{y} - \Psi \hat{\pi}_{WLS}).\end{aligned}$$

The second term $\log(\det(I_p + \gamma^2 \Psi^\top R^{-1} \Psi))$ can be manipulated as follows:

$$\begin{aligned}\log(\det(I_p + \gamma^2 \Psi^\top R^{-1} \Psi)) &= \log(\det(\gamma^2 I_p)) + \log(\det(\gamma^{-2} I_p + \Psi^\top R^{-1} \Psi)) \\ &\approx p \log \gamma^2 + \log(\det(\Psi^\top R^{-1} \Psi))\end{aligned}$$

⁹The symbol \approx denotes equality in the limit as $\gamma \rightarrow \infty$.

where the last approximation clearly holds when $\gamma^2 \rightarrow \infty$. Inserting the last expression in $L_{SPK}(\mathbf{y})$ we obtain that

$$\begin{aligned}L_{SPK}(\mathbf{y}) &\approx \log(\det(2\pi R)) + p \log \gamma^2 + \\ &\quad + \log(\det(\Psi^\top R^{-1} \Psi)) + \\ &\quad + (\mathbf{y} - \Psi \hat{\pi}_{WLS})^\top R^{-1} (\mathbf{y} - \Psi \hat{\pi}_{WLS}) \\ &\approx \hat{L}_{SP}(\mathbf{y}) + \log(\det(\Psi^\top R^{-1} \Psi)) + p \log \gamma^2\end{aligned}$$

which shows that the the two log-likelihoods differ, up to the constant $\log(\det(\gamma^2 I_p)) = p \log \gamma^2$ which is not a function of ρ, σ^2, τ , for a nontrivial term $\log(\det(\Psi^\top R^{-1} \Psi))$ which have an influence on the location of their minima. ■

REFERENCES

- [1] J. J. Craig, *Introduction to robotics: mechanics and control*. Pearson Prentice Hall Upper Saddle River, 2005, vol. 3.
- [2] D. Nguyen-Tuong, M. Seeger, and J. Peters, "Model learning with local Gaussian process regression," *Advanced Robotics*, vol. 23, no. 15, pp. 2015–2034, 2009.
- [3] A. Karimi, M. Butcher, and R. Longchamp, "Model-free precompensator tuning based on the correlation approach," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 5, pp. 1013–1020, Sept 2008.
- [4] M. Fliess, "Model-free control and intelligent PID controllers: towards a possible trivialization of nonlinear control?" *IFAC Proceedings Volumes*, vol. 42, no. 10, pp. 1531–1550, 2009.
- [5] D. Nguyen-Tuong and J. Peters, "Model learning for robot control: a survey," *Cognitive Processing*, vol. 12, no. 4, pp. 319–340, 2011.
- [6] R. Calandra, S. Ivaldi, M. P. Deisenroth, E. Rueckert, and J. Peters, "Learning inverse dynamics models with contacts," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 3186–3191.
- [7] B. Siciliano, L. Sciacivco, L. Villani, and G. Oriolo, *Robotics: modelling, planning and control*. Springer Science & Business Media, 2010.
- [8] J. Hollerbach, W. Khalil, and M. Gautier, "Model identification," in *Springer Handbook of Robotics*. Springer, 2008, pp. 321–344.
- [9] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2006.
- [10] G. Prando, D. Romeres, G. Pillonetto, and A. Chiuso, "Classical vs. Bayesian methods for linear system identification: point estimators and confidence sets," in *Proc. of ECC*, 2016.
- [11] D. Nguyen-Tuong and J. Peters, "Using model knowledge for learning inverse dynamics," in *IEEE International Conference on Robotics and Automation*, 2010.
- [12] T. Wu and J. Movellan, "Semi-parametric Gaussian process for robot system identification," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012, pp. 725–731.
- [13] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [14] B. Bocsi, L. Csató, and J. Peters, "Alignment-based transfer learning for robot models," in *The 2013 International Joint Conference on Neural Networks (IJCNN)*, 2013, pp. 1–7.
- [15] N. Lawrence, M. Seeger, and R. Herbrich, "Fast sparse Gaussian process methods: The informative vector machine," in *Proceedings of the 15th International Conference on Neural Information Processing Systems*, ser. NIPS'02. Cambridge, MA, USA: MIT Press, 2002, pp. 625–632.
- [16] A. J. Smola and P. L. Bartlett, "Sparse greedy Gaussian process regression," in *Advances in Neural Information Processing Systems 13*, T. K. Leen, T. G. Dietterich, and V. Tresp, Eds. MIT Press, 2001, pp. 619–625.
- [17] E. Snelson and Z. Ghahramani, "Sparse gaussian processes using pseudo-inputs," in *Advances in Neural Information Processing Systems*. MIT press, 2006, pp. 1257–1264.
- [18] V. Tresp, "A Bayesian committee machine," *Neural Comput.*, vol. 12, no. 11, pp. 2719–2741, Nov. 2000.
- [19] C. Williams and M. Seeger, "Using the nyström method to speed up kernel machines," in *Advances in Neural Information Processing Systems 13*. MIT Press, 2001, pp. 682–688.
- [20] A. Ranganathan, M. H. Yang, and J. Ho, "Online sparse Gaussian process regression and its applications," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 391–404, Feb 2011.
- [21] L. Csató and M. Opper, "Sparse on-line Gaussian processes," *Neural Comput.*, vol. 14, no. 3, pp. 641–668, Mar. 2002.

- [22] D. Nguyen-Tuong and J. Peters, "Incremental online sparsification for model learning in real-time robot control," *Neurocomputing*, vol. 74, no. 11, pp. 1859–1867, 2011.
- [23] A. Gijsberts and G. Metta, "Incremental learning of robot dynamics using random features," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 951–956.
- [24] A. Rahimi and B. Recht, "Random features for large-scale kernel machines," in *Advances in neural information processing systems*, 2007, pp. 1177–1184.
- [25] R. Camoriano, S. Traversaro, L. Rosasco, G. Metta, and F. Nori, "Incremental semiparametric inverse dynamics learning," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 544–550.
- [26] D. Romeres, M. Zorzi, R. Camoriano, and A. Chiuso, "Online semiparametric learning for inverse dynamics modeling," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 2945–2950.
- [27] J. Quiñero Candela and C. E. Rasmussen, "A unifying view of sparse approximate Gaussian process regression," *The Journal of Machine Learning Research*, vol. 6, pp. 1939–1959, 2005.
- [28] J. Hartikainen and S. Särkkä, "Kalman filtering and smoothing solutions to temporal Gaussian process regression models," in *2010 IEEE International Workshop on Machine Learning for Signal Processing*, Aug 2010, pp. 379–384.
- [29] G. De Nicolao, G. Ferrari-Trecate, and A. Lecchini, "MAXENT priors for stochastic filtering problems," in *Mathematical Theory of Networks and Systems*, Padova, Italy, Jul. 1998.
- [30] M. F. Huber, "Recursive Gaussian process: On-line regression and learning," *Pattern Recognition Letters*, vol. 45, pp. 85 – 91, 2014.
- [31] J. Prüher and M. Simandl, "Gaussian process based recursive system identification," *Journal of Physics: Conference Series*, vol. 570, no. 1, 2014.
- [32] D. Romeres, G. Prando, G. Pilonetto, and A. Chiuso, "Online Bayesian system identification," in *Proc. of ECC*, 2016.
- [33] G. Prando, D. Romeres, and A. Chiuso, "Online identification of time-varying systems: a Bayesian approach," in *Proc. of IEEE CDC*, 2016.
- [34] C. I. Byrnes, A. Isidori, and J. C. Willems, "Passivity, feedback equivalence, and the global stabilization of minimum phase nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 36, no. 11, pp. 1228–1240, Nov 1991.
- [35] J. Taylor, *Classical Mechanics*. University Science Books, 2005.
- [36] G. Pilonetto, M. H. Quang, and A. Chiuso, "A new kernel-based approach for nonlinear system identification," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2825–2840, Dec 2011.
- [37] A. Russu, G. D. Nicolao, I. Poggesi, M. Neve, and R. Gomeni, "Bayesian population approaches to the analysis of dose escalation studies," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 2, pp. 189 – 201, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169260711001519>
- [38] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- [39] G. Wahba, *Spline models for observational data*. Siam, 1990, vol. 59.
- [40] L. Ljung, *System Identification - Theory for the User*, 2nd ed. Upper Saddle River, N.J.: Prentice-Hall, 1999.
- [41] F. Fraccaroli, A. Peruffo, and M. Zorzi, "A new recursive least squares method with multiple forgetting schemes," in *54th IEEE Conference on Decision and Control (CDC)*, 2015, pp. 3367–3372.
- [42] A. Björck, *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, 1996. [Online]. Available: <http://epubs.siam.org/doi/abs/10.1137/1.9781611971484>
- [43] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013, vol. 112.
- [44] K. R. Kozłowski, *Modelling and identification in robotics*. Springer Science & Business Media, 2012.
- [45] M. Zorzi and A. Chiuso, "Sparse plus Low rank Network Identification: A Nonparametric Approach," *Automatica*, vol. 53, 2017.
- [46] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. Von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor *et al.*, "The icub humanoid robot: An open-systems platform for research in cognitive development," *Neural Networks*, vol. 23, no. 8, pp. 1125–1134, 2010.
- [47] F. Janabi-Sharifi, V. Hayward, and C.-S. Chen, "Discrete-time adaptive windowing for velocity estimation," *IEEE Transactions on control systems technology*, vol. 8, no. 6, pp. 1003–1009, 2000.
- [48] S. Ivaldi, M. Fumagalli, M. Randazzo, F. Nori, G. Metta, and G. Sandini, "Computing robot internal/external wrenches by means of inertial, tactile and f/t sensors: theory and implementation on the icub," in *Humanoid Robots (Humanoids), 2011 11th IEEE-RAS International Conference on*, 2011, pp. 521–528.
- [49] U. Pattacini, F. Nori, L. Natale, G. Metta, and G. Sandini, "An experimental evaluation of a novel minimum-jerk cartesian controller for humanoid robots," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1668–1674, 2010.
- [50] F. Nori, S. Traversaro, J. Eljaik, F. Romano, A. Del Prete, and D. Pucci, "icub whole-body control through force regulation on rigid non-coplanar contacts," *Frontiers in Robotics and AI*, p. 18, 2015.
- [51] A. Tacchetti, P. K. Mallapragada, M. Santoro, and L. Rosasco, "GURLS: A least squares library for supervised learning," *Journal of Machine Learning Research*, vol. 14, pp. 3201–3205, 2013.
- [52] J. M. Maciejowski, *Predictive control: with constraints*. Pearson education, 2002.
- [53] G. Pilonetto and A. Chiuso, "Tuning complexity in regularized kernel-based regression and linear system identification: The robustness of the marginal likelihood estimator," *Automatica*, vol. 58, pp. 106 – 117, 2015.
- [54] S. Traversaro, S. Brossette, A. Escande, and F. Nori, "Identification of fully physical consistent inertial parameters using optimization on manifolds," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2016, pp. 5446–5451.