Friction change detection in industrial robot arms

ANDRÉ C. BITTENCOURT

KTH Electrical Engineering

Master's Degree Project
Stockholm, Sweden 2007

XR-EE-RT 2007:026

Abstract

Industrial robots have been used as a key factor to improve productivity, quality and safety in manufacturing. Many tasks can be done by industrial robots and they usually play an important role in the system they are used, a robot stop or malfunction can compromise the whole plant as well as cause personal damages. The reliability of the system is therefore very important.

Nevertheless, the tools available for maintenance of industrial robots are usually based on periodical inspection or a life time table, and do not consider the robot's actual conditions. The use of condition monitoring and fault detection would then improve diagnosis.

The main objective of this thesis is to define a parameter based diagnosis method for industrial robots. In the approach presented here, the friction phenomena is monitored and used to estimate relevant parameters that relate faults in the system. To achieve the task, the work first presents robot and friction models suitable to use in the diagnosis. The models are then identified with several different identification methods, considering the most suitable for the application sought.

In order to gather knowledge about how disturbances and faults affect the friction phenomena, several experiments have been done revealing the main influences and their behavior. Finally, considering the effects caused by faults and disturbances, the models and estimation methods proposed, a fault detection scheme is built in order to detect three kind of behavioral modes of a robot (normal operation, increased friction and high increased friction), which is validated within some real scenarios.

Acknowledgments

This thesis was held at ABB Corporate Research in Västerås. I would like to thank all my colleagues, students and employees, at ABB for the nice moments we shared during my stay in Västerås proportionating a nice work atmosphere. Special thanks for my supervisor, Niclas Sjöstrand, for providing me the opportunity to join the project, who together with the project teammates Johan Gunnar, Shiva Sander-Tavallaey and Sofia Zätterström have shared their knowledge/expertizing and provided an always nice but challenging environment.

I also would like to thank Professor Bo Wahlberg, for keeping me motivated during the thesis and to two special persons at KTH, Cecilia Förssman and Hannelore Eklund for all the support provided.

Finally, this work would have never came true if it was not for the help and support given from good friends and specially from my family. Thank you very much for everything.

Contents

Chapter 1

Introduction

Industrial robots have been used as an important factor to increase productivity and quality in the industry since the past few decades. The first robots to appear have been considered as break-through technologies and represented a huge effort on research and development in each new concept. For several years the efforts on robot development were mostly related to improvements of accuracy and speed of these machines. Nowadays however, industrial robots have reached satisfactory performance levels for what has been its main application field, the manufacturing industry. With the maturity of the robotic systems, new demands appeared to keep the competitiveness which are mostly related to prices reduction and reliability.

The importance of the reliability of a robot is easily understood when taking the example of a robotized assembly line, where the damages caused by an unpredicted stop are counted in function of hours and sometimes minutes. To avoid this situation, it is usual to have some scheduled preventive maintenance of the robot and components in the line. This scheduling, however, is in general based on the estimated robot and components life time and not in its real conditions, remaining a lack of information of the actual system.

The primary objective of this thesis is to define a routine and methods to monitor the mechanical condition of an industrial robot. The approach used to achieve this is the monitoring of the friction phenomena in robot joints. As will be shown, the friction phenomena can relate to faults appearing in the robot and therefore can be used to generate a diagnosis of the system.

1.1 Outline

The thesis outline is as follows.

Chapter 2 presents an introduction to fault detection presenting some remarks to the application sought.

Chapter 3 briefly introduces the field of robotics, presenting the main relevant phenomena and some usual models.

Chapter 4 presents the identification of robot arms where both identification methods and experiments are presented over different models.

In Chapter 5 the friction phenomena in industrial robot arms is presented. Containing a review of the friction phenomena, and a study of the phenomena

in robotics over several operational conditions.

Chapter 6 presents an approach for fault detection based on observations of the friction phenomena in industrial robots. The method is presented and analyzed with some case studies.

Finally, Chapter 7 presents the conclusions of the thesis and leave comments for future work.

In addition, Appendix A presents a more detailed study of the temperature influence on the friction phenomena.

1.2 Contributions

The main contributions of this thesis are:

- The methods and experiments for robot and friction identification presented in Chapter 4.
- The friction behavior of robot joints under several different conditions and variables presented in Chapter 5.
- The fault diagnosis framework based on friction estimated parameters presented in Chapter 6.
- The insights on the temperature influence in Appendix A.

Chapter 2

Fault detection in industrial robots

This chapter presents a review of the main aspects of fault detection and isolation (FDI) methods with the focus on the application on friction change detection in industrial robots.

2.1 Fault detection and isolation overview

This section was based on [3, 13, 18] and provides the reader with a brief overview of FDI (Fault Detection and Isolation) methods. Basically, the purpose of FDI is to monitor dynamic systems and should be able to perform the following tasks:

- *Fault detection*: FDI recognizes that *a fault has occurred*.
- *Fault isolation*: FDI recognizes *where and when* a fault has occurred (some FDI extend this concept to include the type, size or cause of the fault).

When a FDI runs during a normal operation of the system, it is called an *on-line* FDI. If however, the FDI demands the system to be run in a specific manner, the FDI is called *off-line*.

To choose the algorithm of the FDI it is important to know which kind of fault is present in the system. Basically the fault types can be classified by its *time behavior* and *effects on the system*.

The first category of fault can be summarized as:

- *Abrupt*: faults that occur very quickly in the system.
- *Incipient*: faults that occur gradually during time.
- *Intermittent*: faults that affect the system during certain time intervals.

The way a fault affects the system's behavior can be summarized as:

- *Additive*: faults that are effectively added to the system's input or output.
- *Multiplicative*: faults that change the parameters of the system.

- *Structural*: faults that introduce new governing terms to the describing equations of the system.

Figure 2.1 illustrates additive faults in the input signal (f_u) and output signal (f_y) as well as a multiplicative faults (f_{par}) in a system.

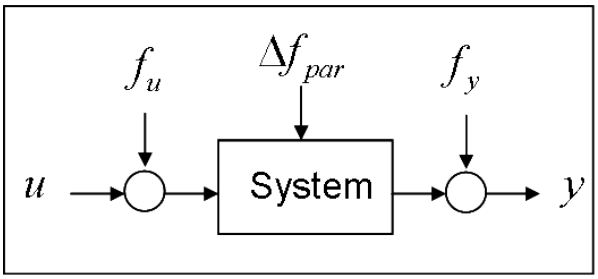


Figure 2.1: Additive and Multiplicative faults

Remark 1 *Friction changes in robotic systems are generally associated with wear and affects a parameter of the system. Therefore, it can be classified as an incipient and multiplicative fault.*

2.1.1 Fault detection and isolation Methods

FDI methods can rely or not on a model of the system. Three convenient categories for *model-free methods* are:

- *Hardware redundancy*: these systems rely on extra hardware which are specially used to detect faults.

- *Spectral analysis*: utilize mechanical vibration, noise, ultrasonic, current or voltage signals to detect and diagnose faults.

- *Expert/Logic systems*: rely on previous knowledge about the behavior and characteristics of the system (age, statistical data, operating condition, etc) under different circumstances. It is a logical method therefore it does not need extra hardware.

Model free methods have been applied successfully in the industry but these methods present some clear drawbacks. In the case of hardware redundancy, extra costs and weight are added to the system; model free methods make use of *a priori* (and often empirical) knowledge of the system signal characteristics, which are dependent on the system operational point and can be costly to define if no previous knowledge about the signals are available.

Dynamic systems like robots have a wide range of operating points making difficult the use of such techniques for FDI, therefore this work will focus on the use of the so-called Model-based methods. The next section gives an overview of the several model-based FDI techniques emphasizing its application to robotic systems.

2.2 Model-based FDI methods

This class of methods are based on the principle of *analytical redundancy*. Instead of comparing several signals outputs for the same variable as in hardware redundancy, they compare analytically generated signals with the system outputs.

Figure 2.2 displays the general flowchart of a model-based FDI method.

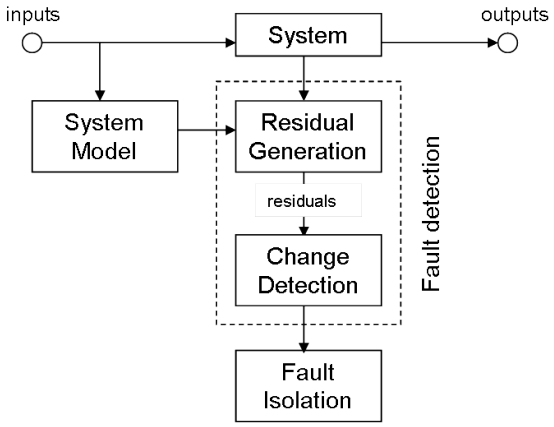


Figure 2.2: Model-based FDI flowchart

Residuals are a fault indicator, based on a deviation between measurements and model-equation-based computations. The residuals are usually generated by filtering techniques that take measured signals and transform them to a sequence of residuals that resemble white noise before a change occurs. There are three main ways of generating residuals in a model-based approach:

- *Parity space*: the system model directly produces outputs that are comparable with the measured outputs.
- *Diagnostic observers or state estimators*: an observer is designed to reconstruct the states of a system, which are compared to the real states generating the residual.
- *Parameter estimation*: an estimation of some physical parameters of the system is compared with its healthy values to generate the residuals.

Since multiplicative faults, like friction change, by definition alter parameters of the system, it makes a natural choice the use of parameter estimation for detection of such faults.

Remark 2 *According to Balle in [4], more than 50% of the applications to detect additive fault use observer methods while more than 50% of the applications to multiplicative faults utilizes parameter estimation methods.*

Nevertheless, one can find several successful applications of observer methods to detect multiplicative faults utilizing for example augmented states where the unknown parameters are modeled as a state of the system. The parity space will not be further discussed since these methods work in an open-loop fashion

which requires a precise model of the system with fixed parameters, which is generally not the case in industrial robots. An example of the use of a parity space approach to fault diagnosis in robots can be found at [22].

In the following subsections, some of the methods for residual generation and fault detection found in the literature will be reviewed and discussed, the focus will be in parameter estimation and diagnostic observers.

2.2.1 Residual generation methods - parameter estimation

Parameter estimation is the process of estimating some parameters of a system model using its input and output measurements. Residuals can be generated when the estimated parameters are compared with fault-free values of such parameters (Figure 2.3). For example, the friction coefficient value can be a good indicator of the condition of a gearbox in a robot joint, monitoring this parameter is then of great interesting for a diagnosis system.

Since the measured signals are stochastic (corrupted by noise) and physical systems are generally nonlinear, recursive estimators like nonlinear observers, extended Kalman filters or recursive least squares are generally used to update the parameter estimates. These parameters are usually initially guessed and then converge to a final value after multiple recursive steps.

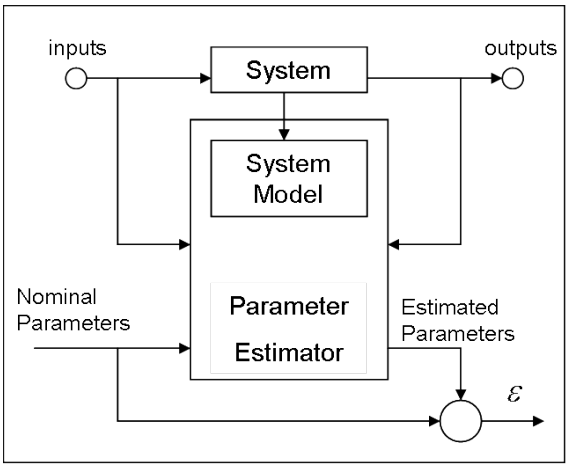


Figure 2.3: Parameter Estimation Block Diagram

There are various different but related conceptual bases for continuous-time system parameter estimation, see [3, 21]. They are briefly described here.

1. *Output error methods, OE*: this is maybe the most intuitive parameter estimation approach. The parameters are estimated in order to minimize the error between the model output and the system output.

$$\varepsilon'(t) = y(t) - \frac{\hat{B}}{\hat{A}}u(t) \tag{2.1}$$

where $\hat{A}$ and $\hat{B}$ are the estimates of the governing polynomials of the system. In this case, no direct calculation of the parameters is possible, because $\varepsilon(t)$ is nonlinear in the parameters. The loss function is therefore minimized as an optimization problem.

2. *Equation error methods, EE*: this approach is clearly derived from an analogy with static regression analysis and linear least squares estimation. The error function is generated directly from the input-output equations of the model.

$$\varepsilon'(t) = \hat{A}y(t) - \hat{B}u(t) \tag{2.2}$$

From Equation (2.2) it is clearly seen that it implies the generation of the time derivatives of the signal, which might be a problem when the signal is too noisy. Young [21] proposes a solution for this utilizing a 'generalized equation error' that filters the measured signals and provides filtered derivatives of the signals.

After sampling, the estimation can be solved as a least square estimate or in a recursive form (recursive least squares). Isermann, see [3], emphasizes that for numerical properties improvement, square-root filters algorithms are recommended.

3. *Prediction error methods, PE*: the equation of the error is the same as the OE case, Equation (2.1), the difference is that the output estimate is defined as a 'best prediction' depending on the current estimates of the parameters a which characterize the system and the noise models, $\hat{y}(t)\widehat{=}\hat{y}(t|a)$. $\hat{y}(t|a)$ is a conditional estimate of $y(t)$ given all current and past information of the system, while $\varepsilon(t)$ is an 'innovations' process with serially uncorrelated white noise characteristics (see [31] for more). The PE can be written as an OE, Equation(2.3), or EE, Equation (2.4).

$$\varepsilon'(t) = \frac{\hat{C}}{\hat{D}}\left[y(t) - \frac{\hat{B}}{\hat{A}}u(t)\right] \tag{2.3}$$

$$\varepsilon'(t) = \frac{\hat{C}}{\hat{D}\hat{A}}\left[\hat{A}y(t) - \hat{B}u(t)\right] \tag{2.4}$$

4. *Maximum likelihood methods, ML*: a special case of PE methods, separated here because of its importance, with the additional restriction that the stochastic disturbances to the system have specified amplitude probability distribution functions. In several applications, this assumption is restricted further for analytical tractability to the case of a Gaussian distribution.

5. *Bayesian methods*: extension of ML where *a priori* information on the probability distributions is included in the formulation of the problem. It is important in the FDI context because most recursive methods can be interpreted as being a Bayesian type.

An usual solution for parameter estimation is the use of linear models. Here we summarize the main parameter estimation methods for *linear* continuous-time models based on sampled signals, see [7] for more.

1. *Least-squares parameter estimation*: this is a well-known case of optimization where the estimated parameters vector $\hat{\theta}$ are estimated by the non recursive estimation equation below.

$$\hat{\theta} = [\psi^T \psi]^{-1} \psi^T y \tag{2.5}$$

   where ψ is the data vector and y is the measured output. These parameters are biased by any noise, therefore, a good signal-to-noise ratio must be achieved to use this method.

2. *Determination of the time derivatives*: as mentioned before, the estimation of the signals derivatives by numerical differentiation is not a good approach because of the inherent noise in the signals. A state variable filter is therefore utilized that calculates the derivatives and filter the noise.

3. *Instrumental variables parameter estimation*: instrumental variables can be used to overcome the bias problem due to noise. The instrumental variables introduced are only insignificantly correlated with the noise-free process output. A major advantage of instrumental variables is that no strong assumptions and knowledge on the noise is required. However, when dealing with closed loop configurations biased estimates are obtained because the input signal is correlated with the noise.

4. *Parameter estimation via discrete-time models*: one can try to estimate the variables in discrete-time models and then calculate the parameters of the continuous-time model. These methods, however, require extensive computational effort and are not so straightforward.

2.2.2 Residual generation methods - state estimation (observers)

This category of FDI methods uses a state observer to reconstruct the unmeasurable state variables based on the measured inputs and outputs. It can be shown that an additive fault is easily detected with this technique, this kind of fault makes the residual (generally taken as the estimation error) deviate from zero with a bias.

Remark 3 *The influence of multiplicative faults in residuals generated by state observers is not as straightforward recognizable because in this case the changes in the residuals could be caused either by parameter, input and state variable changes.*

The main advantage of observer-based methods is that they do not require special excitation of the system, making it a *good choice for on-line fault detection.*

Observer-based FDI methods also require an accurate mathematical model of the process, therefore it is important to try to robustify the residual evaluation in order to cope with the inherited uncertainties of any physical model.

Investigations of robust observer-based approach can be found for example at [17, 15].

The following FDI methods with state estimation are known, see[3]:

1. *Dedicated observers for multi-output processes*: the design of specific observers allows the detection of specific faults, combining and arranging the observers one can detect multiple faults.
   (a) *Observer excited by one input*: one observer is driven by one sensor output while the other outputs are estimated and compared with the measures allowing the detection of single sensor faults (additive faults).
   (b) *Bank of observers, excited by single outputs*: several of the first case allowing the detection of multiple sensor faults.
   (c) *Kalman filter, excited by all outputs*: the residuum changes the characteristic of zero mean white noise with known covariance if a fault appears, which is detected by a hypothesis test.
   (d) *Bank of observers, excited by all outputs*: several of the above designed to detect a definite fault signal.
   (e) *Bank of observers, excited by all outputs except one*: as before, but each observer is excited by all outputs except one sensor output which is supervised.
2. *Fault detection filters for multi-output processes*: the feedback state observer is chosen so that particular fault signals in the input change in a definite direction and fault signals at the output change in a specific plane.
3. *Output observers*: another possibility is the use of output observers (unknown input observers) if the reconstruction of the state variable is not of primary interest. A linear transformation is applied so that the residuals are dependent only on additive input/output faults.

2.3 Change (fault) detection methods

After the generation of the residuals, it is needed to establish whether there was a change (fault) on the system or not. This role is done by the change detector which can be classified under three categories, see [18]:

1. *One model approach*: The filter residuals ε_t are transformed to a distance measure s_t (computed from the no-fault values), a stopping rule decide whether the change is relevant or not. A schematic is show at Figure 2.4.

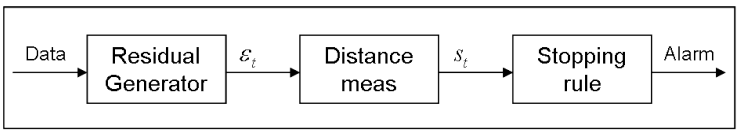


Figure 2.4: One model approach for change detection

The most natural distance measures are:

- Change in the mean, $s_t = \varepsilon_t$.
- Change in the variance, $s_t = \varepsilon_t^2 - \lambda$, where λ is a known fault-free variance.
- Change in correlation, $s_t = \varepsilon_t y_{t-k}$ or $s_t = \varepsilon_t u_{t-k}$ for some k.
- Change in sign correlation, $s_t = sign(\varepsilon_t \varepsilon_{t-1})$, this test is used due to the fact that white residuals should change sign every second sample in the average.

2. *Two model approach*: In this case the residuals are generated by two filters, a slow (with a great data window or the whole data) and a fast one (with a small data window) which are compared, Figure 2.5 illustrates the procedure. If the model based on the smaller data window gives larger residuals, than a change is detected. The main problem is to choose an adequate norm for the comparison, typical norms are:

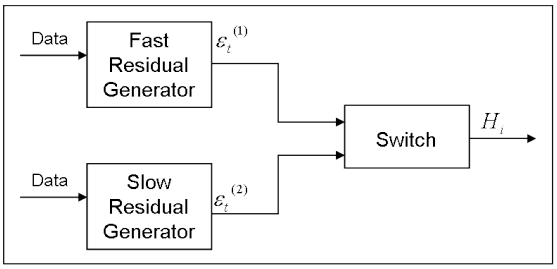


Figure 2.5: Two residual generators running in parallel, one slow to get good noise attenuation and other fast to get fast tracking. The switch decides whether a change occurred or not.

- The Generalized Likelihood Ratio (correlation between fault signatures).
- The divergence test.
- Change in spectral distance.

Remark 4 *The choice of the window size of the fast filter is a trade-off between quick detection and accurate model (avoiding false alarms).*

3. *Multi-model approach*: This approach makes use of the so-called *matched filters*, that can generate white residuals for a specific change even after it was inserted in the system. The idea is to enumerate all conceivable hyphoteses about changes and compare the residuals generated from the matched filters, the one with the 'smallest' residuals will be an indication of the change, Figure 2.6 shows the procedure. Since a batch of data is needed, this approach is off-line, but many proposed algorithms makes the calculations recursively, and are consequently on-line.

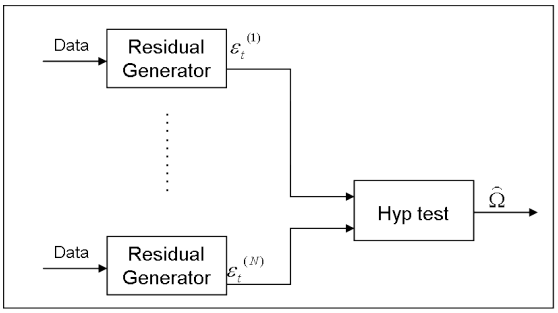


Figure 2.6: Several matched filters (residual generators) that are compared in a hyphotesis test.

2.4 Concluding remarks

The chapter presented a review of some fault detection methods. Emphasis has been given on methods for monitoring unmeasurable quantities like process parameters and process state variables.

In designing of FDI methods, the following aspects should be taken in consideration, see [7]:

- *Process models*: since the methods are based on the deviation of a normal operation, one should define the normal operation of the system (for example, nominal values of parameters) and also which kind of model. If the system or process is running only with small changes of the variables, linearized models can be used. However for many applications this is not the case (see for example [16, 14, 23]) and one should take this in consideration while defining which kind of model should be used. Isermann and Ballé, see [4], mention that there is an increase on the use of non-linear models for parameter estimation.

  Besides the use of analytical models (change detection of the outputs of an analytical model), a diagnosis system can also rely on heuristics of the system. These heuristics can be translated for example in fault-symptom-trees or fuzzy logic and are important for the fault isolation.

- *Parameter and state estimation*: as discussed through the chapter, state estimation has its main applications in the detection of additive faults and has the drawback that it is difficult to identify the source of the fault since the residuals are deviations of the system states.

  On the other hand, parameter estimation techniques are the most indicated approach for the identification of multiplicative faults. On [6, 26] several multiplicative faults could be identified utilizing parameter estimation, validating its importance. A main drawback of parameter estimation is that the system input signal must be informative enough to identify the parameters of the system, which makes it difficult to implement in on-line diagnosis systems.

  In a complete diagnosis system, where both additive and multiplicative faults are required, the FDI methods utilizing parameter and state estimation complements each other. For example, an observer-based FDI

method that detects faults on sensor and actuators could run on-line, whether an actuator fault is detected an off-line test utilizing a parameter estimation FDI method could be used to fault diagnosis and isolation.

- *Faults*: the way that a fault affects the system is very important when designing FDI methods (see Section 2.1). It is also important to define the main faults present in the system.

- *Performance*: fault detections must be sensitive to the appearance of faults but insensitive to other changes (noise, operating points, modeling errors, etc.). Because these requirements often contradict each other, the following trade-offs must be analyzed:

  - size of fault *vs* detection time;
  - speed of fault appearance *vs* detection time;
  - speed of fault appearance *vs* process response time;
  - size and speed of fault *vs* speed of process parameters changes;
  - detection time *vs* false alarm rate.

  Methods that are sensitive to abrupt faults for example, might not be suitable to detect incipient faults. Therefore, several methods can be used in parallel.

- *Practical aspects*: a FDI method should take in consideration the practical aspects when defining the experiments to detect the faults. An industrial robot application generally contains several restrictions (restricted work envelope of the robot, measurements with noise, limited computational effort, limited sensors available, system under feedback action, etc.) and should be robust enough to cope with them.

- *Testing*: the introduction of artificial faults is also important to validate the system reliability and should try to approximate to real faults.

Chapter 3

Robotics and robot modeling

This chapter presents a general description of an industrial robot and dynamic models considering its relevant aspects in the context of a parameter based fault detection such as nonlinearities caused by flexibilities and backlash.

3.1 Industrial robots

An industrial robot (Figure 3.1) can be described as a mechanical manipulator which is programmable and controlled to achieve tasks as moving objects and tools through a predetermined trajectory. The mechanical structure of a standard industrial robot is composed by *links* and *joints*. Links are the main bodies that make up the mechanism, these links are connected in pairs by joints. The way the links are connected by the joints define the kinematic chain of a robot, if one link is only connected to one other link, the robot is called serial. According to the application a tool is also coupled at the output link of the robot.

A joint add constrains to the relative movement of the links connected to it. According to these constrains, a joint can be called for example revolute (permits rotation in one direction between the links), prismatic (allows linear movement in one direction), etc, see [27]. According to the kind of joints and its number, a robot will have more or less d*egrees of freedom* (DOF). Here, serial robots with 6 revolute joints will be considered (6 DOF) . The first three joints gives mobility to the arm and the other three (also called wrist) proportionate orientation to the end-effector of the robot. Nowadays, industrial robot joints are driven in general by electric motors and a gearbox to give the necessary torque to move the links.

An industrial robot is a complex system where the dynamics and kinematics aspects are very important to its general performance. Therefore, there is a demand for realistic models of robots to use in simulation, control design and diagnosis. However, it is not an easy task to design an accurate global model due to a lot of non-linearities and phenomena that are not fully understood. The following sections presents some usual models of industrial robots and the main phenomena present.

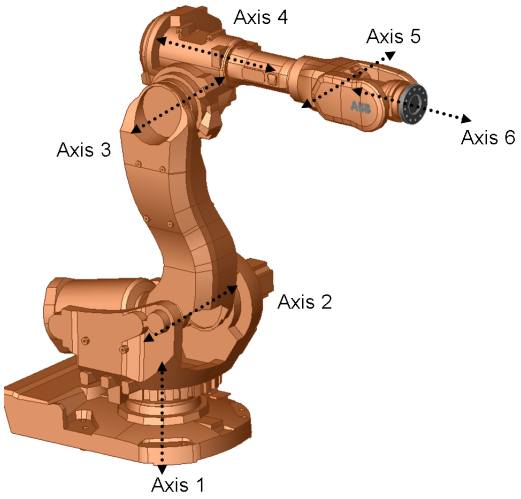


Figure 3.1: An ABB industrial robot (IRB 6600) and its axes

3.2 Robot modeling overview

There are some known aspects of a robot that should be considered while designing a model:

Flexibilities In general, robot links can be considered as rigid-bodies (no flexibilities) but nowadays there is a special demand on reducing production costs which generally means reducing weight. Making lighter links reduces the production costs of the mechanical part of a robot (which represents more than 50% of its total price) but on the other hand, the flexible modes of the links get more evident. Also, the gearboxes present in a joint, specially the harmonic drive type, introduces flexibilities due to elastic deformation of bearings and gears. Such flexibilities are nonlinear which makes the system more challenging to control, model and estimate.

Friction Friction affects any mechanical moving parts and has extensively been studied due to its importance in mechanical systems. Even though there is no analytical model for friction, there are some well-known friction phenomena and models based on empirical experiments to describe it. An usual way of modeling friction is to consider only its static effects like Coulomb, viscous and Stribeck friction (see Chapter 5 for more aspects of the friction phenomena).

Backlash Backlash is present in all mechanical system where the motor is not directly coupled to the load. It can be described as the clearance between mating components when movement is reversed and contact is re-established. For a gear for example, the backlash is the amount of clearance between mated gear teeth. When the backlash gap is opened, the movement of the load is autonomous and the moment generated by the motors drives only the motor itself and not the load. Several attempts to control, model and identify backlash can be found in the literature, see [28, 29] for examples.

Torque Ripple In general, electric AC permanent magnet motors are used as joint actuators. This motors are compact, fast and robust. A drawback is that the generated torque changes periodically with the rotor position. Distortion of the stator flux linkage distribution and variable magnetic reluctance at the stator slots are the main causes of the resulting torque ripple. The ripple caused by the magnetic reluctance is proportional to the current and periodical in the rotor position and affects the performance of the system. See [24, 30] for more.

Measurement inaccuracies Besides the inherent noise in any measurement system, position measurements in robots are generally obtained by using Tracking Resolver-to-digital converters, which error can be modeled as a sum of sinusoids.

The choice of a model is always a trade-off between fidelity of the model (how well it represents the real system) and its complexity (inclusion of nonlinearities, flexible modes, etc). Each application requires a level of fidelity of the model, for example, if the model should be used in an accurate simulation, one should try to include all physical aspects to the model.

In the next sections, some dynamic models are proposed to accomplish the task of parameter based fault detection in industrial robots. The models of the arm presented in Sections 3.3 and 3.4 are derived under the simplifications that:

- the controllers of the electrical motors are neglected and the torque reference is viewed as the applied torque to the system. This assumption is fair since the electrical dynamics of the motors are comparatively much faster than the mechanical system.
- the arm is modeled in an axis without the influence of gravity (axis one for example). The aspects of the gravity influence will be discussed in Section 3.5.
- the backlash is neglected. Backlash models will be discussed in section 3.6.
- only one axis is excited at a time.[1]

The notation presented in Table 3.1 is used through the chapter.

3.3 Rigid body model

The first model presented is a classic *two-mass rigid-body* model of a robot arm as show in Figure 3.2

The model is composed by two masses, J_m representing the joint inertia and J_a representing the links (arm) inertia. The shafts inertia and backlash are not included and both joints and arm are considered as rigid-bodies. The masses are coupled through an ideal gearbox with ratio r. The assumption of a stiff coupling between the masses gives $\varphi_a = r\varphi_m$. The representative dynamic

[1] Coupling forces are then neglected.

Table 3.1: Notation used

Parameter	Description
J_x	inertia at x side. When only J is used it relates to the whole robot inertia.
r	gearbox ratio
φ_x	position at x side.
τ_x	torque at x side. When only τ is used it relates to the applied motor torque.
f_c	Coulomb friction parameter.
f_v	viscous friction parameter.
d_x	damping constant at x side.
k_x	stiffness constant at x side.
CoG	center of gravity.
φ_{CoG}	angle formed from moving axis and center of gravity.
M	arm mass.
g	gravity constant.

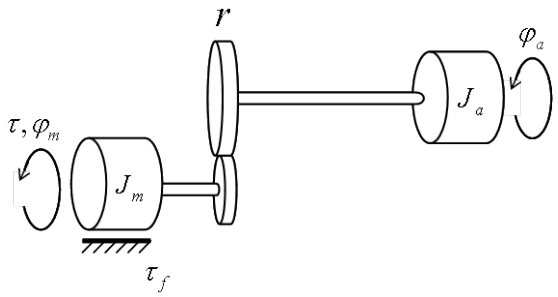


Figure 3.2: The two mass rigid body model

equations are:

$$\begin{aligned} \tau &= J_m\ddot{\varphi_m} + J_a\ddot{\varphi_a} + \tau_f \\ \tau &= (J_m + r^2 J_a)\ddot{\varphi_m} + \tau_f \end{aligned} \tag{3.1}$$

3.4 Including flexibilities

3.4.1 Two mass flexible model

A two-mass flexible model can be utilized to include all flexibilities in one spring coupling two masses, Figure 3.3.

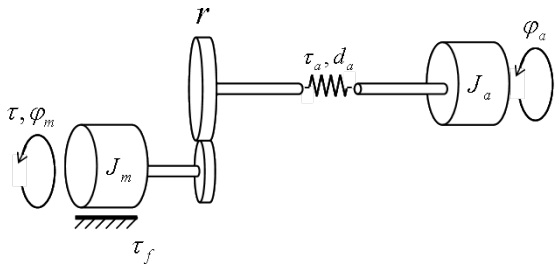


Figure 3.3: Two mass flexible model

The resulting dynamic equations for this model are:

$$\begin{aligned} J_m\ddot{\varphi_m} + rd_a(r\dot{\varphi_m} - \dot{\varphi_a}) + \tau_f + r\tau_a &= \tau \\ J_a\ddot{\varphi_a} - d_a(r\dot{\varphi_m} - \dot{\varphi_a}) - \tau_a &= 0 \end{aligned} \tag{3.2}$$

Where τ_f is the friction torque and τ_a the spring torque. The spring torque can be represented by a simple linear model with one parameter $\tau_a = k_a(r\varphi_m - \varphi_a)$ or as a nonlinear function $\tau_{a_{NL}} = k_{a_1}(r\varphi_m - \varphi_a) + k_{a_2}(r\varphi_m - \varphi_a)^3$.

The two mass flexible model can be considered as a good representation of the manipulator when it is not moving too fast and the arm can be considered stiff when the only relevant flexibilities are related to the gearbox.

3.4.2 Three mass flexible model

To separate the gearbox and arm flexibilities, the model is extended to a three mass model.

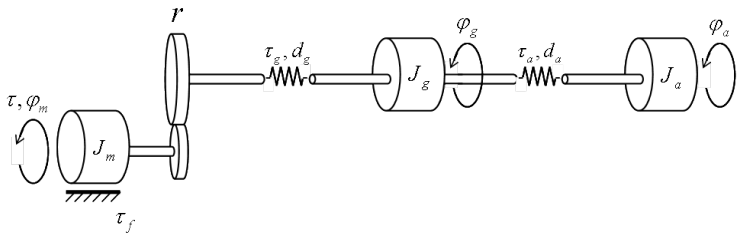


Figure 3.4: Three mass flexible model

The masses in Figure 3.4 represents the motor, gearbox and arm from left to right. The model dynamic equations are:

$$\begin{aligned} J_m\ddot{\varphi_m} + rd_g(r\dot{\varphi_m} - \dot{\varphi_g}) + \tau_f + r\tau_g &= \tau \\ J_g\ddot{\varphi_g} + d_a(r\dot{\varphi_g} - \dot{\varphi_a}) - d_g(r\dot{\varphi_m} - \dot{\varphi_g}) + \tau_a - \tau_g &= 0 \\ J_a\ddot{\varphi_a} - d_a(r\dot{\varphi_g} - \dot{\varphi_a}) - \tau_a &= 0 \end{aligned} \tag{3.3}$$

See [34] for more on flexibilities modeling and identification.

3.5 Including gravitational forces

In axes where the position of the moving joint does not change the resulting gravitational force (for example axis 1 at Figure 3.1) it is only dependent on the position of the other joints and the load. For axes where these torques are dependent on the moving joint position, like axis 2, a simple model can be proposed.

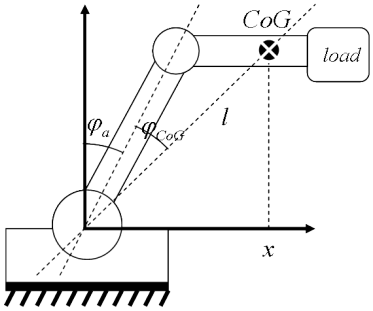


Figure 3.5: Gravitational torque in a robot arm

Figure 3.5 models the gravitational torque acting in axis 2 of a robot arm. According to the robot and load masses distribution, the Center of Gravity (CoG) around axis 2 will be somewhere in the space. φ_a and φ_{CoG} are angles in the arm side. The resulting torque on the joint will be

$$\begin{aligned} \tau_{CoG} &= Mgx \\ \tau_{CoG} &= Mglsin\left(\varphi_a + \varphi_{CoG}\right) \\ \tau_{CoG} &= Mgl\left(sin(\varphi_a)cos(\varphi_{CoG}) + cos\varphi_a sin(\varphi_{CoG})\right) \end{aligned} \tag{3.4}$$

The resulting gravitational torque is then added to the model and acts in the applied torque as

$$\tau' = \tau - \tau_{CoG} \tag{3.5}$$

Where τ' is the new applied torque to the joint. Considering the movement in only one axis at a time φ_{CoG} remains constant, otherwise φ_{CoG} will be a function of the distribution of the links and load masses in the space, which may be difficult to estimate, specially if the load mass is unknown.

The gravitational torque modeled in this section assumes that the robot center of motion is in the same plane of its center of gravity. If this is not the case a component of the gravitational force will also appear along the axis that is moving. Nowadays there is a trend to design asymmetric robots to achieve bend-over movements which makes the force along the axis more evident.

3.6 Backlash models

Figure 3.6 is an extension of the rigid-body model presented in Section 3.3 including backlash in the shaft.

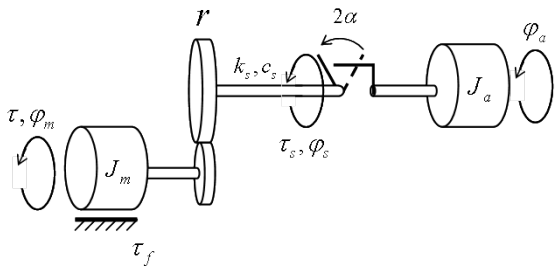


Figure 3.6: Backlash included in the rigid body model

A classical approach to model backlash is to consider it as a deadzone, neglecting damping, where the shaft torque is proportional to the shaft twist, $\theta_s = k_s D_\alpha(\theta_d)$ where $\theta_d = \theta_m - \theta_a$ is the displacement angle between motor and arm and the deadzone function ($D(x)$)

$$D(x) = \begin{cases} x - \alpha & x > \alpha \\ 0 & \mid x \mid < \alpha \\ x + \alpha & x < \alpha \end{cases} \tag{3.6}$$

Including the damping we have the shaft torque as

$$\theta_s = k_s \theta_s + c_s \dot{\theta_s} \tag{3.7}$$

defining the backlash angle $\theta_b = \theta_d - \theta_s$ it is possible to obtain the dynamic equation

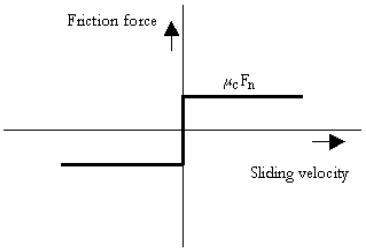
$$\theta_b = \begin{cases} max\left(0, \dot{\theta_d} + \frac{k_s}{c_s}(\theta_d - \theta_b)\right) & \theta_b = -\alpha(\tau_s \leq 0) \\ \dot{\theta_d} + \frac{k_s}{c_s}(\theta_d - \theta_b) & \mid \theta_b \mid < \alpha \\ min\left(0, \dot{\theta_d} + \frac{k_s}{c_s}(\theta_d - \theta_b)\right) & \theta_b = \alpha(\tau_s \geq 0) \end{cases} \tag{3.8}$$

3.7 Friction models

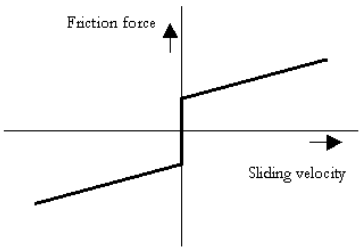
There are several models for friction proposed in the literature, see for example [1, 10, 11, 12]. A brief summary of these models and its properties are presented here.

Classical models Whether there is no need for the use of a high fidelity model of the friction, one can model it as a simple static model. Figure 3.7 shows usual examples of static friction models.

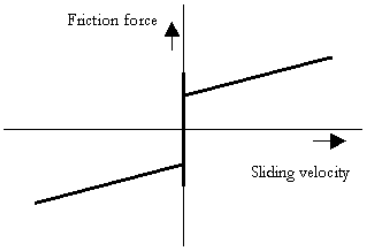
It is important to note that this kind of models are not causal, since the discontinuity at zero speed allows the friction to assume several values. Some solutions for this problem can be found at [1, 11].



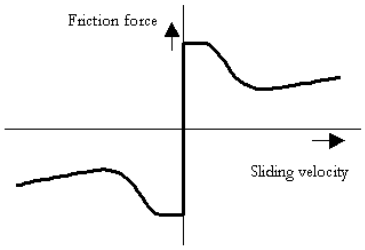
(a) Coulomb friction



(b) Coulomb and Viscous friction



(c) Static, Coulomb and Viscous friction



(d) Static, Coulomb, Viscous and Stribeck friction

Figure 3.7: Static friction models. 3.7(a), the friction component that is only dependent of the direction of velocity, not of the magnitude of the velocity; 3.7(b), the friction component that is proportional to velocity (viscous) and goes to zero at zero velocity with the Coulomb term; 3.7(c), the torque or force necessary to initiate motion from rest (the so called break-away force, generally larger than the Coulomb term) with the static and viscous terms; 3.7(d), the friction phenomenon that arises from the use of fluid lubrication and gives rise to decreasing friction with increasing velocity at low velocity with Coulomb, static and viscous terms.

Dynamic models Nowadays, the interest in dynamic models for friction has increased due to demands for precision servos as well as advances in hardware that makes feasible the implementation of friction compensators. Several dynamic models have been purposed in the literature, see for example [1, 10]. Here two of them are presented:

- *Bliman-Sorine* The Bliman-Sorine model is a second order model (4 parameters) that can be seem as a parallel connection of two Dahl models, see [1]. It models static, viscous and Coulomb friction of the static friction phenomena and only pre-sliding displacement of the dynamic friction phenomena.
- *LuGre* The LuGre model can be seen as a first order Dahl model (6 parameters) with a velocity-varying coefficient to give stiction. The model is inspired by the bristle interpretation of friction in combination with lubricant effects. It models the Stribeck effect and also the rate dependent friction phenomena such as varying break-away force and frictional lag.

	Coulomb Model	Coulumb and Viscous Model	Coulomb, Viscous and Static Model	Coulomb, Viscous, Static and Stribeck Model	Dahl Model	Bliman-Sorine Model	LuGre Model
Static Friction Phenomena							
Coulumb Friction	X	X	X	X	X	X	X
Viscous Friction		X	X	X		X	X
Static Friction			X	X		X	X
Stribeck Friction				X			X
Dynamic Friction Phenomena							
Pre-sliding displacement					X	X	X
Varying break-away force							X
Frictional Lag							X
Number of parameters	1	2	3	≥ 4	2	4	6

Table 3.2: Comparison of friction models

Table 3.2 shows that the LuGre Model is the most complete one. The drawback is that it also demands the estimation of more parameters. For most engineering applications though, a static friction model is enough, experimental works comproved that a good static friction model can approximate real friction forces with 90% of confiability, see [11]. Some efforts on using more complete models can be found in [1, 12].

The simple model including the Coulomb and viscous terms can be described by:

$$\tau_f = f_c sign(\dot{\varphi_m}) + f_v \dot{\varphi_m} \tag{3.9}$$

Where f_c represents the direction-dependent *Coulomb friction* and f_v is the velocity proportional *viscous friction*. A great advantage of this model is its simplicity with only two parameters.

The more realistic model, presented by Feeny-Moon in [2], describe also the nonlinearities of the friction phenomena present in the low velocities:

$$\tau_f = f_v \dot{\varphi_m} + f_c \left(\mu + \frac{(1-\mu)}{cosh(\beta \dot{\varphi_m})} \right) \tag{3.10}$$

The friction phenomena in robot joints will be further investigated in Chapter 5.

3.8 Concluding remarks

This chapter introduced the robotics field presenting the main phenomena and robot models. The choice of a model depends on the application, in the context of fault detection using parameter estimation for example, the most important is the parameter capability to relate faults. Not always a more complex model will give more information about the faults.

The next chapter presents the results of the identification of industrial robot arms focusing the task of friction change detection.

Chapter 4

Robot identification

As mentioned by Wernholt in [25], parameter identification in Robotics can be divided in at least three groups: robot kinematics, robot dynamics (often divided in rigid body and flexible body models) and joint model. The kinematics parameters are generally obtained through CAD softwares while the other two are usually experimentally identified.

Robot dynamics identification covers inertial (rigid body) as well as flexibilities parameters (elastic effects on the robot structure). Joint identification involves motor inertia, friction, backlash, and gearbox flexibilities.

This chapter covers the identification of robot dynamics and joint parameters. First a method to identify the rigid body parameters and joint friction is presented and analyzed in detail followed by robot identification including joint flexibilities, finally Section 4.4 presents the identification of friction parameters over its characteristic curve.

Before continuing it is important to take some considerations about parameter identification in industrial robots:

- *Feedback influence*: sometimes it is necessary to operate the identification under the influence of feedback (*closed-loop*). This is the case of industrial robots where feedback is needed to maintain the arm in a desired position. Two challenges arises when identifying systems under feedback influence. First, it will be a non-zero correlation between the input signal and the disturbance of the measured output. The second is that the data contain less information about the open-loop system since the purpose of feedback is to make the closed-loop less sensitive to changes in the open-loop system. *Prediction error methods* works well for systems under feedback if the model represents the true system and the data is informative enough.

- *Restricted input signal*: the controller architecture of commercial industrial robots is usually closed, with only point-to-point programming available under pre-defined velocities. The use of more complex signals like multi-sines and chirps are then restricted.

- *Restricted sensors available*: in industrial robots, usually only variables at the motor side are measured, and it includes motor applied torque, motor position and motor velocity, meaning that without including extra sensors, the variables at the arm side are not available.

These properties certainly influences the identification method and experiments used and should be considered.

4.1 Rigid body parameters estimation

As presented in Chapter 3, the dynamic model of a rigid body arm under gravity influence can be represented by the following equation:

$$\begin{aligned} \tau &= J_m\ddot{\varphi_m} + J_a\ddot{\varphi_a} + \tau_f + \tau_{CoG} \\ \tau &= (J_m + r^2 J_a)\ddot{\varphi_m} + \tau_f + \tau_{CoG} \end{aligned} \tag{4.1}$$

where τ, τ_f and τ_{CoG} are respectively motor applied torque, friction torque and gravity torque.

A simple representation of friction can include the Coulomb and viscous terms:

$$\tau_f = f_c sign(\dot{\varphi_m}) + f_v\dot{\varphi_m} \tag{4.2}$$

and the gravitational torque:

$$\tau_{CoG} = Mglsin(\varphi_a)cos(\varphi_{CoG}) + cos\varphi_a sin(\varphi_{CoG}) \tag{4.3}$$

leading to

$$\begin{aligned} (J_m + r^2 J_a)\ddot{\varphi_m} + f_c sign(\dot{\varphi_m}) + f_v\dot{\varphi_m} \\ +Mgl(sin(\varphi_a)cos(\varphi_{CoG}) + cos\varphi_a sin(\varphi_{CoG})) = \tau \end{aligned} \tag{4.4}$$

Considering that commanded torque to the motor (τ), motor velocity ($\dot{\varphi_m}$) and acceleration ($\ddot{\varphi_m}$) can be measured or estimated, Equation (4.4) leaves five parameters to be identified: the robot inertia, $J = (J_m + r^2 J_a)$, the velocity proportional friction coefficient, f_v, the direction of movement proportional friction coefficient, f_c and the gravitational terms $Mglcos(C\hat{o}G)$, $Mglsin(C\hat{o}G)$.

In this manner Equation (4.4) can be rewritten as the linear regression:

$$\widehat{\tau_k} = \Phi_k\Theta^T + e_k \tag{4.5}$$

$$\begin{aligned} \Phi_k &= [\ddot{\varphi_{m,k}} \quad \dot{\varphi_{m,k}} \quad sign(\dot{\varphi_{m,k}}) \quad sin(\varphi_{a,k}) \quad cos(\varphi_{a,k})] \\ \Theta &= [J \quad f_v \quad f_c \quad Mglcos(\varphi_{CoG}) \quad Mglsin(\varphi_{CoG})] \end{aligned}$$

Θ, $\widehat{\tau_k}$, Φ_k and e_k are respectively the parameters vector, predicted motor torque, data vector and noise sampled terms.

Defining the *prediction error* as the difference between the measured motor torque, $\tau(k)$, and the predicted motor torque $\widehat{\tau}(k|\theta)$

$$\varepsilon(k,\theta) = \tau(k) - \widehat{\tau}(k|\theta) \tag{4.6}$$

and choosing a minizing criteria for Equation (4.6) such as the quadratic prediction error

$$V_N(\theta) = \frac{1}{N}\sum_{k=1}^{N}\frac{1}{2}\varepsilon^2(k,\theta) \tag{4.7}$$

Since the prediction error here is linear in the parameters, the minizing vector $\hat{\theta}_N$ for V_N is the solution to a standard least-squares problem

$$\hat{\theta}_N = arg \overset{\theta}{min}\, V_N(\theta) = \left[\frac{1}{N}\sum_{k=1}^{N}\Phi(k)\Phi^T(k)\right]^{-1} \frac{1}{N}\sum_{k=1}^{N}\Phi(k)\tau(k) \tag{4.8}$$

which is easy to compute and analyze. For more information on system identification and parameter estimation see Ljung [31].

4.1.1 Experiment design

The concept of *informative experiments* defines that a data set Z is informative enough with respect to a model set M^*, meaning that the data allow discrimination between any two different models in M^*. Basically this can be translated as requirements to the input signal to the system. The main objective is to use an input signal that *excites all relevant frequencies* for the identification and has a small *crest factor*, which is a measure of the distribution of the input power at the frequencies range (signals with small crest factor have a spread amplitude spectrum).

Scope The experiments were held with the robot configuration as an inverted L, only axis 1 was moved while the other axes remained still. Even though the identification has been performed in all axis, for simplification reasons only axis 1 (without gravity influence) will be considered here, the results, however, can easily be extended to other axes.

Input signal choice In order to identify the rigid body parameters, with focus on the friction parameters, it is desirable that the input signal (here the motor velocity is considered as input) attend the following requirements:

- *Low frequencies of excitation*: as will be shown in Section 4.2, the fricton parameters are affected only by low frequency range signals therefore it is desirable that the input signal amplitude is concentrated in the low frequency range.

- *Steady state velocities*: if the acceleration is too high, the flexible modes of the system will be excited and therefore the assumption of a rigid body model is not valid and the estimation will be biased.

The objective now is to design an input signal under the limitations mentioned that is informative enough to perform identification and that only needs to be activated for a short time.

The chosen signal is shown in Figure 4.1. It was generated by changing the velocity of the robot as a stair between two points ($-50°$ to $50°$). The signal is activated for 40 seconds and goes through 20 different steady-state velocities.

Data preprocessing The system is excited while commanded torque and velocity are acquired. No acceleration measurements are available, which is estimated from the velocities measurements by the *central difference algorithm*:

$$\ddot{\varphi}(k) = \frac{\dot{\varphi}(k+1) - \dot{\varphi}(k-1)}{2T_s} \tag{4.9}$$

In order to improve the SNR (Signal to Noise Ratio) of the estimated acceleration a non-causal lowpass filter is used.

To generate the $sign(\dot{\varphi_m})$ signal avoiding fluctuations near zero velocity caused by the high frequency noise, first the velocity measurements are lowpass filtered and then an adapted function of $sign(x)$ is used such as:

$$sign\_thold(x, thold) = \begin{cases} 0 & \text{if } |x| < thold \\ sign(x) & else \end{cases} \tag{4.10}$$

Before estimating the model it is good to have a measure of the data set quality to the estimation. The concept of **condition number** can be used for this purpose. The condition number associated with the linear equation $Ax = b$ gives a bound on how inaccurate the solution x will be after approximate solution. For this problem the condition number can be defined as

$$\begin{aligned} cond\_number(\Phi_k) &= ||{\Phi_k}^{-1}||.||\Phi_k|| \\ &\quad \text{for } ||\Phi_k||_2 \\ cond\_number(\Phi_k) &= \frac{\sigma_{max}(\Phi_k)}{\sigma_{min}(\Phi_k)} \end{aligned} \tag{4.11}$$

where $\sigma_{max}(\Phi_k)$ and $\sigma_{min}(\Phi_k)$ are the maximal and minimal singular values of Φ_k. A problem with small condition number is said to be *well-conditioned*, while a problem with a high condition number is said to be *ill-conditioned*.

Estimation The final data set used in the estimation, with condition number 400.47, can be seen in Figure 4.1.

The estimated parameters values with its confidence intervals are shown in Table 4.1.

Parameter	Identified Values
J	0.0184 ± 0.0002
f_c	0.6042 ± 0.0849
F_v	0.0107 ± 0.0007

Table 4.1: Estimated parameters

4.1.2 Validation

The next step is to validate the model to assure that the system is well represented by the model. A new data set is used which was generated by exciting the same trajectory as in the estimation but in reverse way.

Figure 4.2(a) shows a model fit higher than 70% which is a reasonable result considering the simplified model used for both robot dynamics and friction. The cross-correlation values, Figure 4.3, are inside the confidence interval (the auto-correlation is high because of the feedback present in the system).

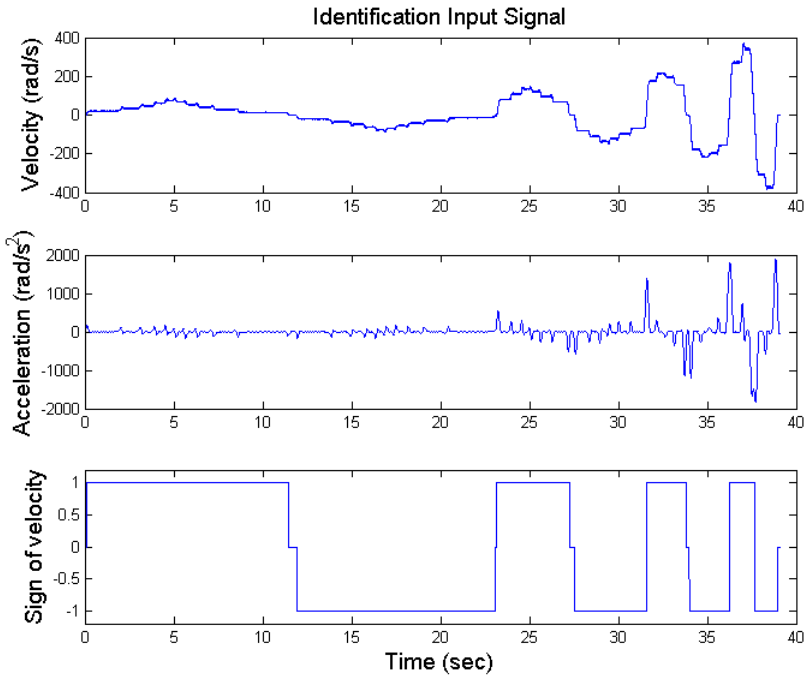
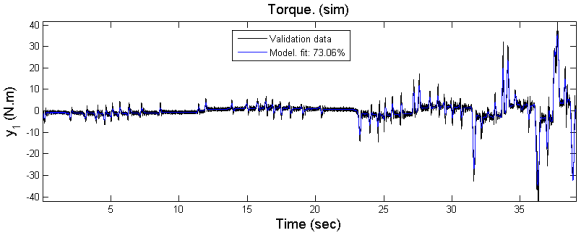
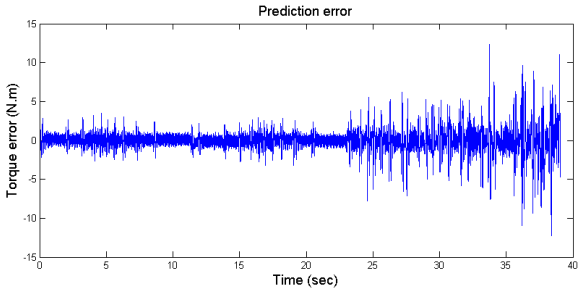


Figure 4.1: Identification Input Signal, note the higher acceleration values for the higher velocities.

Figure 4.2(b) shows the prediction error, Equation (4.6), between the predicted and measured torque, where it is easy to realize that the error is greater for higher acceleration values, which is explained by the simplified model used.



(a) Model fit to validation data



(b) Model prediction error.

Figure 4.2: Model validation

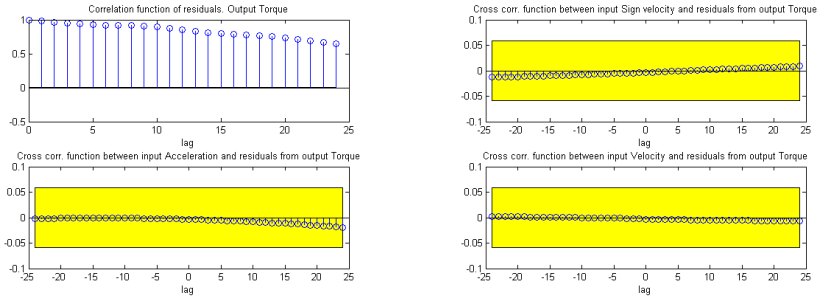


Figure 4.3: Correlation Analysis.

4.1.3 Evaluation of RB parameters identification method

It is evident that the simplified rigid body model will not be a good representation of the system for all kind of inputs and operational conditions. Therefore, further analysis on the influence of the acceleration of the input signal and on the use of simplified trajectories have been performed.

Input Acceleration Influence

As discussed in Section 4.1.1 an important issue when choosing the input signal to the identification of rigid body is the avoidance of high accelerations. As seem in Figure 4.1, the acceleration assumes high values for this trajectory, specially for the higher speeds. To check the acceleration influence on the estimated parameters as well as on the numerical conditioning of the problem, the data used on Section 4.1.1 was filtered by Algorithm 5, which takes away the data from the data vector Φ where the acceleration values are greater than a threshold:

Algorithm 5 *Acceleration thresholding*

j=1;
for i=1:K
  if $|\ddot{\varphi}_m(i)| \leq thold$
    then
      $\Phi_{NEW}(j) = \Phi(i);$
      $\tau_{NEW}(j) = \tau(i);$
      $j = j + 1;$
  end
end

where K is the length of the data vector Φ, Φ_{NEW} and τ_{NEW} are the new filtered data vector and applied motor torque respectively and *thold* is the acceleration threshold of the filter. With Φ_{NEW} and τ_{NEW}, the linear regressor from Equation (4.6) is used to estimate the parameters through a range of different *thold* values. The result can be seem in Figure 4.4.

Remark 6 *This approach is only valid because the linear regressor from Equation (4.8) at sample k only depends of the data sampled at k and not from past or future data.*

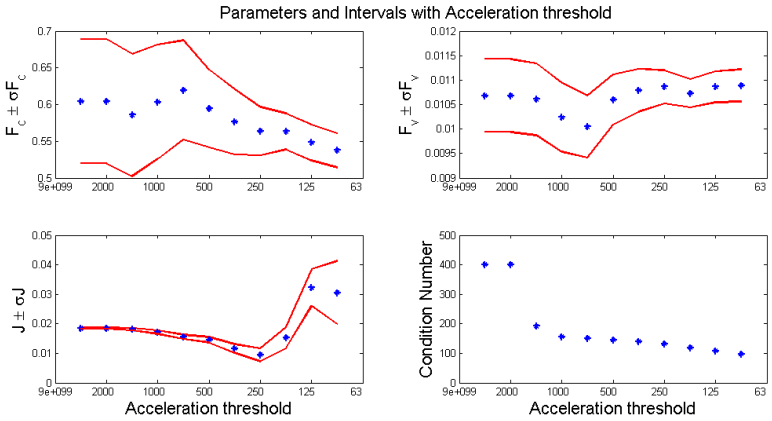


Figure 4.4: Parameters with its standard deviation (limiting lines) and the condition number for different acceleration threshold values.

The following conclusions can be drawn about the effect of thresholding the acceleration on the estimation signal.

- $\downarrow thold \Rightarrow \downarrow \sigma f_v$ *and* σf_c: this result confirms that the friction parameters are more affected by the low frequency input as already discussed. However, it is only considerable when the threshold is smaller than 200 (10% of the maximum acceleration observed).

- $\downarrow thold \Rightarrow \uparrow \sigma J$: this result is consistent since the inertia parameter is directly determined by the acceleration signal.

- $\downarrow thold \Rightarrow \downarrow$ *Condition number*: the decrease on the condition number is specially relevant in the first iterations, where the condition number could be reduced by more than half. Of course that this conclusion is not valid for too low thresholds which would reduce the data size too much and increase the condition number.

Simpler trajectories

To verify the assumption of a rigid body model and the estimation method presented in Section 4.1.1, the method was performed with simpler identification data sets. The identification sets were generated moving the robot axis between two points $[-30°, 30°]$ with different commanded speeds. The parameters and condition number of each data vector Φ are then compared.

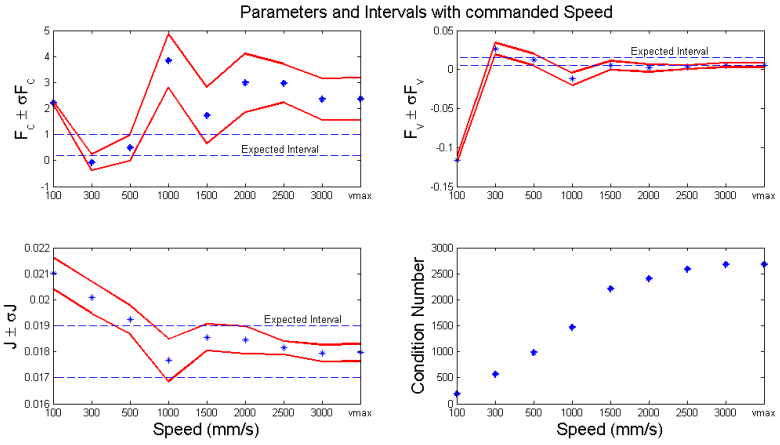


Figure 4.5: Rigid body parameter estimation applied to point-to-point trajectories with different speeds.

Analyzing the results at Figure 4.5 it is easy to conclude that the rigid body identification method will only produce good results when the input signal is informative enough. The condition number increases exponentially which indicates that this approach is not valid, *restricting the use of this method to well conditioned input signals*. For better results with such simple trajectories, more complex models can be used and\or more variables measured.

4.1.4 Concluding remarks

Some characteristics of the rigid body identification method proposed:

- *The estimation quality is directly dependent on the input signal.* The method will only produce good results if the input signal is constrained so that it only excites the desired frequencies for the identification and passes through several steady state velocities. This characteristic restricts the method application to situations where the arm can be excited in a certain range of the work envelope so that an informative enough trajectory can be applied. This also restricts its use to off-line identification using a specific trajectory to the identification.
- *The method is simple and fast*, the data can be rapidly collected and processed and the identification is done in only one step.

4.2 Joint flexibilities parameters identification

One of the main sources of flexibilities in industrial robots are in the gearboxes which appear due to elastic deformation of bearings and gears. As presented in Chapter 3 a model for the robot arm including flexibilities can be represented by following dynamic equations

$$\begin{aligned} J_m\ddot{\varphi_m} + r_a d_a(r\dot{\varphi_m} - \dot{\varphi_a}) + \tau_f + r\tau_a &= \tau \\ J_a\ddot{\varphi_a} - d_a(r\dot{\varphi_m} - \dot{\varphi_a}) - \tau_a &= 0 \end{aligned} \tag{4.12}$$

where the subscripts $_a$ relates the variables to the arm side (after the gearbox) and $_m$ to the motor side. A challenge that arises with this model is that the $_a$ variables cannot be measured or estimated and therefore, the problem cannot be solved by simple linear regressors and the use of black box techniques are then needed.

In this document a state space linear model of Equation (4.12) is used to estimate the parameters. Considering the linear models for τ_f and τ_a as

$$\begin{aligned} \tau_f &= f_v\dot{\varphi_m} \\ \tau_a &= k_1(r\varphi_m - \varphi_a) \end{aligned}$$

the corresponding state-space model in the form of $\dot{x} = Ax + Bu$, $y = Cx + Du$ is

$$X = \begin{bmatrix} r\varphi_m - \varphi_a & \dot{\varphi_m} & \dot{\varphi_a} \end{bmatrix}^T \tag{4.13}$$

$$\begin{aligned} A &= \begin{bmatrix} 0 & r & -1 \\ \frac{-(rk_1)}{Jm} & -\frac{(f_v+r^2d)}{Jm} & \frac{rd}{Jm} \\ \frac{k_1}{Ja} & \frac{rd}{Ja} & -\frac{d}{Ja} \end{bmatrix} \\ B &= \begin{bmatrix} 0 & \frac{1}{J_m} & 0 \end{bmatrix}^T \\ C &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \\ D &= \begin{bmatrix} 0 \end{bmatrix} \end{aligned} \tag{4.14}$$

with τ as input and $\dot{\varphi_m}$ as output. The model can be considered as a state space representation with known structure and unknown elements.

4.2.1 Experiment design

Before estimating the model it is important to check at which frequencies each of the unknown parameters are excited. This is easily seem by looking at the frequency response of the state-space model described by Equation (4.14) when the parameters are changed. Figure 4.6 shows the results for a 20% of changes in the nominal values of the parameters (Table 4.2).

Parameter	Nominal value
f_v	0.0107
k_1	1500
d	4000
J_m	0.0043
J_a	709

Table 4.2: Nominal values. The nominal value for J_m has been taken from the manufacturer specification while f_v and J_a from the rigid body identification step ($J = (J_m + r^2 J_a)$). k_1 and d were guessed.

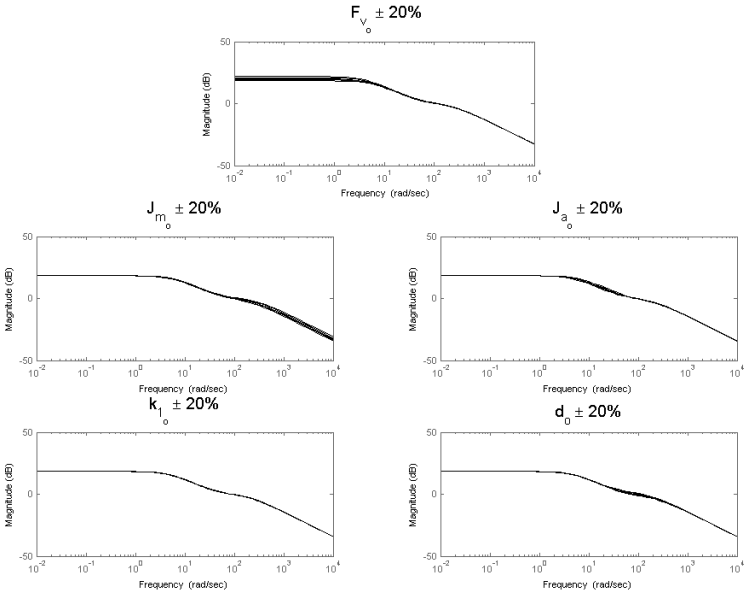


Figure 4.6: Parameters exciting frequencies.

The following conclusions can be drawn from Figure 4.6:

- f_v affects the low range frequencies.
- J_m affects the high range frequencies.

- d and J_a affects the middle range frequencies.
- k_1 also affects the middle ranges but with only a small influence.

Input signal choice Two main considerations might be taken when designing the input signal for the identification of the two mass model with joint flexibilities of Equation (4.14):

- *Excite all range of frequencies*: as seem above, the parameters excites all range of frequencies.
- *Avoidance of static friction*: since the linear model do not include the static friction f_c term, it is advisable that the input signal have as less as possible zero velocities crossings.

The signal chosen for the estimation is a periodic triangle wave as the one shown in Figure 4.7.

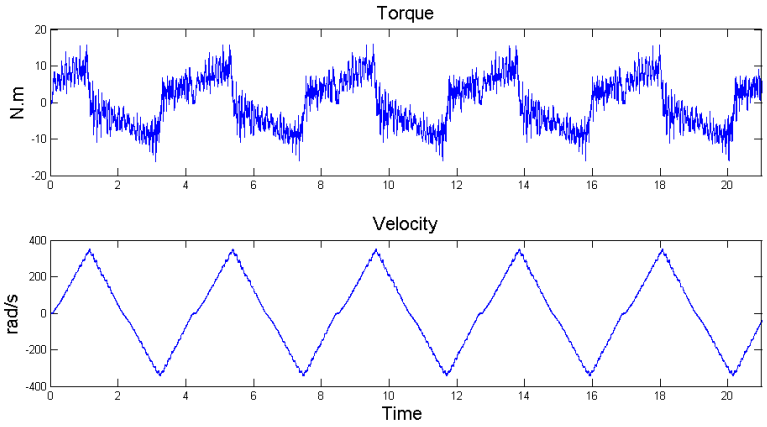


Figure 4.7: Periodic estimation input signal.

Data preprocessing To improve the SNR of the estimation signal, it has been averaged over the periods with an arithmetic average. This also reduces the amount of data in which the model will be estimated and consequently the processing time.

Estimation Given the initial values for the parameters as the ones found in Table 4.2, the model is estimated using an iterative prediction error method. Function `pem` from the MATLAB© *System Identification Toolbox* is used to estimate the state-space model of the form

$$\begin{aligned} x(t+Ts) &= Ax(t) + Bu(t) + Ke(t) \\ y(t) &= Cx(t) + Du(t) + e(t) \end{aligned} \tag{4.15}$$

where K is the noise model which was set to zero during the estimation.

Data post processing The estimated model returns a set of parameters on the state space form such as:

$$\begin{aligned} A &= \begin{bmatrix} 0 & r & -1 \\ \theta_1 & \theta_2 & \theta_3 \\ \theta_4 & \theta_5 & \theta_6 \end{bmatrix} \\ B &= \begin{bmatrix} 0 & \theta_7 & 0 \end{bmatrix}^T \end{aligned}$$

to obtain the physical parameters from the estimated black box model, the linear equations are solved as

$$\begin{aligned} J_m &= \frac{1}{\theta_7} \\ k_1 &= \frac{-\theta_1 J_m}{r} \\ J_a &= \frac{k_1}{\theta_4} \\ d &= \frac{\theta_5 J_a}{r} \\ f_v &= -\theta_2 J_m - r^2 d \end{aligned} \tag{4.16}$$

The estimated parameters obtained after 1000 iterations are shown in Table 4.3.

Parameter	Estimated value
f_v	0.0157
$k1$	2881.3
d	52982
J_m	0.0088
J_a	59.6364

Table 4.3: Estimated values.

4.2.2 Validation

To validate the model the data set used in the identification of the rigid body parameters is used.

Besides the evident improvements with this model there is still a bias for velocities near zero as seem in Figure 4.8. This behavior can be explained by the non included static friction f_c in the model.

4.2.3 Conclusions

Some characteristics of the rigid body identification method proposed:

- *Sensibility to initial values* of the estimated model. Since the identification is proceeded in a black box model the choice of the initial parameter values are very important, the results can vary a lot with the initial conditions, sometimes converging to unfeasible values (negative friction or inertia).

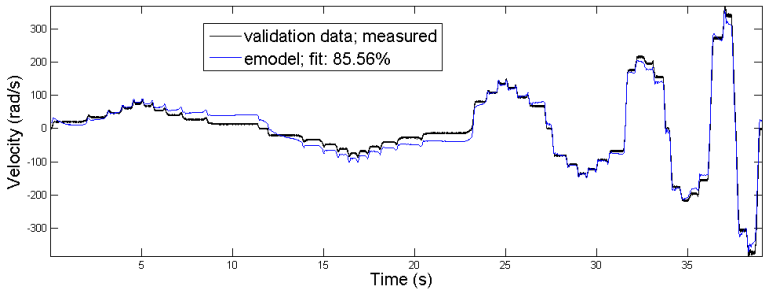


Figure 4.8: Validation of the linear state space model with data used on the rigid body identification.

It is very important then that the rigid body parameters are consistent before estimating the model.

- *Non inclusion of static friction* produces a bias on the estimates for velocities near zero.

4.3 Non-linear grey box identification

Given the initial parameters as the result of the other two presented estimations methods, a non-linear model which includes the static friction is used to achieve a better performance. The model is defined by the dynamic equations:

$$\begin{aligned} \dot{x_1} &= rx_1 - x_2 \\ \dot{x_2} &= \frac{-r(k_1 x_1) - rd(rx_2 - x_3) - f_v x_2 - f_c sign(x_2) + \tau}{J_m} \\ \dot{x_3} &= \frac{k_1 x_1 + d(rx_2 - x_3)}{J_a} \end{aligned} \tag{4.17}$$

where

$$\begin{aligned} x_1 &= r\varphi_m - \varphi_a \\ x_2 &= \dot{\varphi_m} \\ x_3 &= \dot{\varphi_a} \end{aligned} \tag{4.18}$$

the initial parameters values are set as the ones identified in the steps described before which values are shown in Table 4.4.

The input signal for estimation is the same as the one used in Section 4.2.1, a triangle wave averaged over periods. f_c and f_v are set as fixed parameters and the others set as free. The model is estimated using an iterative prediction error method, function `pem` from the MATLAB© *System Identification Toolbox*. The estimated parameters are shown in Table 4.5.

4.3.1 Validation

Three data sets have been used to evaluate the final model:

Parameter	Nominal values
f_c	0.6042
f_v	0.0107
k_1	2881.3
d	52982
J_m	0.0043
J_a	59.6364

Table 4.4: Nominal values.

Parameter	Estimated value
f_c	0.6042
f_v	0.0107
k_1	2769.0
d	51544
J_m	0.013
J_a	203.0

Table 4.5: Estimated values.

1. Validation data 1: same data used to validate the rigid body model (Figure 4.9).

2. Validation data 2: same data used to validate the joint flexibilities model (Figure 4.10).

3. Validation data 3: point-to-point trajectory with one predetermined velocity (Figure 4.11).

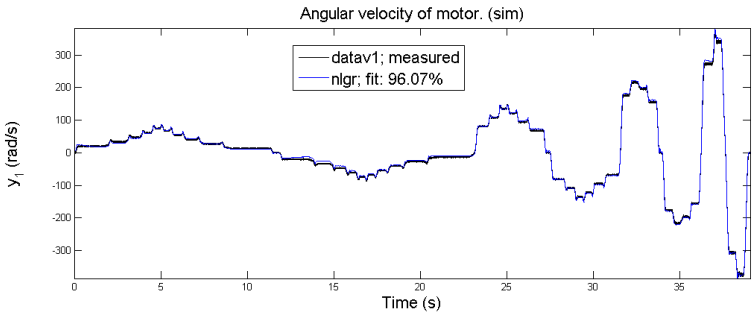


Figure 4.9: Validation data 1 FIT.

4.3.2 Conclusions

The non linear grey box model presents the best performance amongst all proposed models, reducing the bias for low velocities and representing the system well for several types of inputs. The drawback is that, as the linear flexible model, it is also sensible to the initial values, besides of being more complex, requiring longer estimation times.

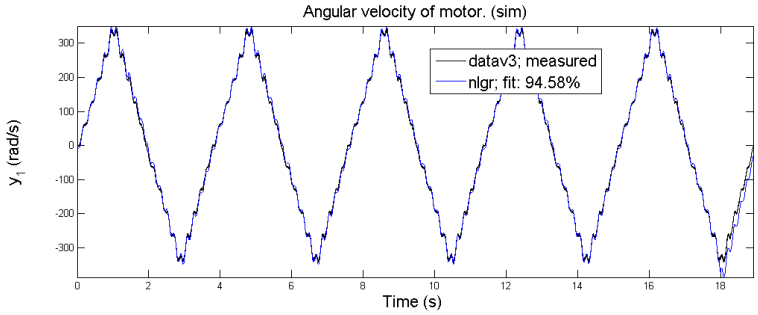


Figure 4.10: Validation data 2 FIT.

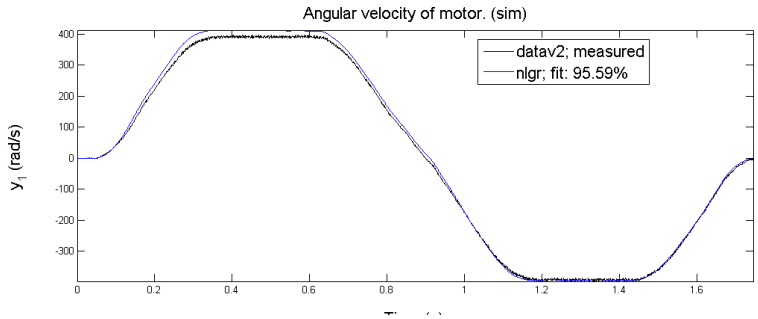


Figure 4.11: Validation data 3 FIT.

4.4 Friction parameters identification

The identification methods presented until now in this chapter are used to describe the robot system completely including the identification of dynamics and joint parameters. This section, however, deals only with the identification of the friction phenomena.

Besides of utilizing data from the system measured signals, the friction parameters are estimated from the friction characteristic curve. First, then, a method to estimate the friction curve is presented.

4.4.1 Friction curve estimation method

Equation (4.19) presents a generalized model of an industrial robot arm, where $M(\ddot{\varphi_a})$ is the inertia matrix, $C(\varphi_a, \dot{\varphi_a})\dot{\varphi_a}$ is referred as to the velocity dependent term and includes the centrifugal, Coriolis and friction forces and $\tau_g(\varphi_a)$ is the torque that appears due to the gravity and is position dependent (changes on the center of mass). τ is the applied torque to the robot (in the motor side).

$$M(\ddot{\varphi_a}) + C(\varphi_a, \dot{\varphi_a})\dot{\varphi_a} + \tau_g(\varphi_a) = \tau \tag{4.19}$$

Supposing that *only one joint is moved at each time*. For steady-state velocities the trajectory does not excite flexible modes so that $\varphi_a = r\varphi_m$ and also Coriolis and centrifugal forces can be neglected, Equation (4.19) is then reduced to:

$$\tau_f(\dot{\varphi_m}) + \tau_g(\varphi_a) = \tau \tag{4.20}$$

where $\tau_f(\dot{\varphi_m})$ is the friction torque.

Neglecting the directional dependency of the friction phenomena, the gravitational torque $\tau_g(\overline{\varphi_a})$ at the validation point $\overline{\varphi_a}$ and friction torque $\tau_f(\overline{\dot{\varphi_m}})$ at velocity $\overline{\dot{\varphi_m}}$ can be estimated as:

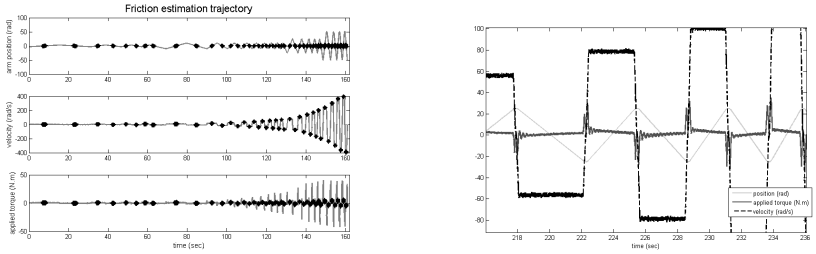
$$\begin{aligned} \tau_{fwd} &= \tau_f(\overline{\dot{\varphi_m}})_{fwd} + \tau_g(\overline{\varphi_a}) \\ \tau_{bwd} &= -\tau_f(\overline{\dot{\varphi_m}})_{bwd} + \tau_g(\overline{\varphi_a}) \\ \tau_g(\overline{\varphi_a}) &= \frac{\tau_{fwd} + \tau_{fwd}}{2} \\ \tau_f(\overline{\dot{\varphi_m}}) &= \frac{\tau_{fwd} - \tau_{fwd}}{2} \end{aligned} \tag{4.21}$$

Position, velocity and applied torque measurements in the motor side are usually available in robot applications. Therefore, using the method presented, one can estimate the value of the friction torque for several different validation velocities $\overline{\dot{\varphi_m}}$ and draw a curve for friction torque *vs* velocity.

Experiment design

The objective now is to define a trajectory that excites the system through several different steady-state velocities and an algorithm to compute the friction torque at each steady-state velocity.

The trajectory proposed (gray curve in Figure 4.12(a))goes through 22 different steady-state velocities and is generated with simple point-to-point trajectories for each different velocities.



(a) The identification trajectory is marked in gray while the estimation data in black.

(b) Gravity influence. Note the slope of the applied torque in steady-state velocities.

Figure 4.12: Identification trajectory and gravity influence.

An algorithm searches for the steady-state velocities preparing the data for the estimation. Two main considerations might be taken when designing such algorithm.

Gravity influence For axes under *gravity influence*, axis 2 for example, the applied torque at a steady-state velocity is not constant since the gravitational torque will be a function of the arm position (see Chapter 3 for more). Therefore, the gravitational torque influence requires that the data used in the estimation to be taken at the same position $\overline{\varphi_a}$ in the backward and forward directions for each steady-state velocity (see Figure 4.12(b)).

Noises A basic requirement is that the evaluation torque and velocity values are estimated as precise as possible besides the noise. In the method presented here, this is achieved by taking the average of torque and velocity measurements around the validation point $\overline{\varphi_a}$ inside an interval $(\overline{\varphi_a} - lim, \overline{\varphi_a} + lim)$.

The final data used for the estimation of friction is marked in black at Figure 4.12(a).

The data is then processed as presented earlier in this section to take away the gravity influence and estimate friction. The result of such estimation is a curve with the friction torque at several velocities as shown in Figure 4.13.

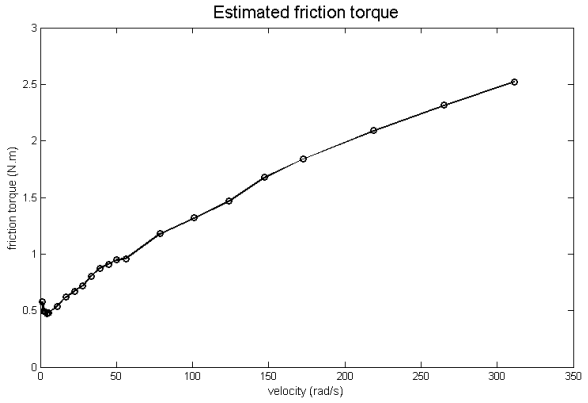


Figure 4.13: Estimated friction torque (N.m).

4.4.2 f_c and f_v estimation

The simple model that includes only the viscous and Coulomb friction parameters can be estimated over the friction characteristic curve with the simple linear regressor below:

$$\begin{aligned} \widehat{\tau_f} &= \Phi_k \Theta^T + e_k \\ \\ \Phi_k &= [\dot{\varphi_{m,k}} \qquad sign(\dot{\varphi_{m,k}})] \\ \Theta &= [f_v \qquad f_c] \end{aligned}$$

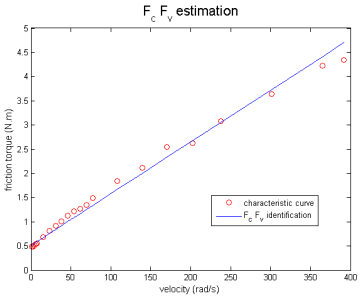
The result of the estimation is presented together with real friction curves in Figure 4.4.2.

From Figure 4.14(b) it is easy to see that the simple model that approximates all the friction phenomena for non zero velocities as a straight line can give big errors.

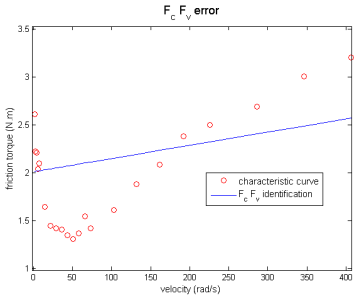
4.4.3 A more complete model

The more realistic friction model

$$\tau_f = f_v \dot{\varphi_m} + f_c \left(\mu + \frac{(1-\mu)}{cosh(\beta \dot{\varphi_m})} \right) \tag{4.22}$$



(a) Characteristic curve with little nonlinearities.



(b) Characteristic curve with high nonlinearities.

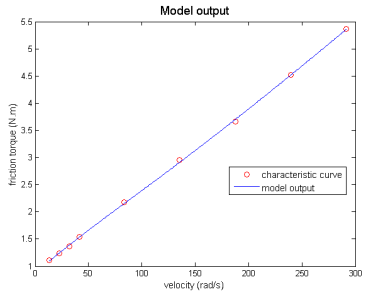
Figure 4.14: Friction estimated model output and friction characteristic curves.

gives a better estimation of the friction phenomena for the whole velocities range. The model, however, is nonlinear in the parameters and cannot be estimated directly with a simple linear regression.

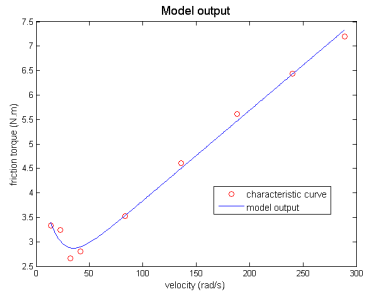
To estimate the model parameters, the linear regression is combined with extensive search that varies the parameters μ and β in a predefined range until a prediction error equation is minimized. The linear regressor is updated for each pair of values μ and β as:

$$\begin{aligned} \widehat{\tau_f} &= \Phi_k \Theta^T + e_k \\ \Phi_k &= \left[\dot{\varphi_{m,k}} \qquad \left(\mu + \frac{(1-\mu)}{cosh(\beta \dot{\varphi_m})} \right) \right] \\ \Theta &= [f_v \qquad f_c] \end{aligned}$$

From the model output presented in Figure 4.4.3 it is easy to realize the better representation of this model.



(a) Characteristic curve with little nonlinearities.



(b) Characteristic curve with high nonlinearities.

Figure 4.15: Friction estimated model output and friction characteristic curves.

4.5 Concluding remarks

This chapter covered dynamics and joint parameters identification of industrial robots. The choice of which model to use and the experiment to identify is one of the most important tasks in parameter driven fault detection and can be considered as a trade off between reliability of the parameter and the complexity of the identification method.

Table 4.6 presents a summary of each of the methods presented here.

Model	Technique	Parameters	Advantages	Disadvantages
Rigid body	linear regressor	f_c, f_v, J	Simple method and simple experiment (one step identification). Parameters identified directly from the data.	Poor representation of the real system.
Joint flexibilities	Black box state-space	f_v, J_a, J_m, $k1$, d	Good representation of the system.	Two step identification. sensibility of the estimation to the initial parameters choice.
Joint flexibilities	Non linear Grey Box	f_c, f_v, J_a, J_m, $k1$, d	Best representation of the system.	Three step identification. Time consuming estimation. Sensibility to initial parameters choice.
Simple friction	linear regressor from characteristic curve.	f_c, f_v	Simple method and experiment. Parameters identified directly from the characteristic curve.	Poor representation of the characteristic curve.
Complete friction	linear regressor with ext search from charac curve.	f_c, f_v, μ, β	Simple method and experiment. Parameters identified directly from the characteristic curve.	Extensive search can be time consuming.

Table 4.6: Identification methods comparison.

Chapter 5

Friction phenomena in robot joints

Friction exists in all mechanism to some extent. As defined by Olsson in [1], friction is the tangential reaction force between two surfaces in contact, which is physically dependent on contact geometry and topology, properties of the bulk and surface materials of the bodies, displacement and relative velocity of the bodies and presence of lubrication.

The success of a fault detection system is related to its ability to generate alarms only when a real faulty state is present in the system, avoiding false alarms. Therefore, it is vital to have as much knowledge as possible about how the monitored variables are influenced by the system. In this sense, this chapter presents the friction behavior under several different circumstances, delimiting the boarders of a fault detection using friction as a parameter.

5.1 Friction phenomena

Friction is a highly nonlinear phenomena which has been constantly explored by researchers due to its importance in a lot of applications, including the robotics field. Since friction is the result of complex interactions between contact surfaces and lubricants, there is not a predictive model for the phenomena based only in the characteristics of the materials. Therefore, the models are invariably defined through experimental data.

Generally, friction has been modeled with simple static models (constant velocities), but nowadays, the interesting on dynamic models for friction has increased considerably in order to have some better estimation of the friction.

Some of these dynamical behaviors of friction can be summarized below:

- *Pre-sliding displacement (Dahl effect)* The spring-like behavior of friction that causes a displacement linear dependent on the applied force if this applied force is less than the break-away force (force necessary to overcome the static friction).

- *Varying break-away force (rising static friction)* The dependence of the break-away force on the rate of increase of the applied force.

- *Frictional Lag* The delay in the change of the friction force as a function of a change in the velocity.

Besides the friction phenomena mentioned above, some other effects with respect to friction are reported in the literature:

- *Time-dependent friction* From experiments it is known that friction changes with time. These changes appears, for instance, due to loss of lubricant, deformation of the surface material, change in temperature due to generated heat and/or accumulation of wear debris.

- *Position-dependent friction* A dependence of the friction on the position of a system is another effect that is experimentally observed and is well known by a lot of researchers. This position-dependency is caused by spatial inhomogeneities in the transmission of the system due to contact geometry and/or loading which varies as a function of position. As the load varies, the normal force between the sliding surfaces varies, causing a varying friction (friction is linear dependent on the normal force). By preloading the transmission elements and roller bearings this dependency of friction on the load can be decreased.

- *Direction-dependent friction* A lot of researchers have found the friction to be dependent on the direction of the motion of a system. Different Coulomb and viscous friction levels in the left and right directions of a single, linear motion have been observed experimentally on many occasions. Theoretically, this may be due to anisotropies in material or geometry.

The friction phenomena is usually classified as pre-sliding and sliding, as illustrated in Figure 5.1. The Coulomb friction, or the phenomena present in low velocities (pre-sliding range) can be explained by the complex interactions between the surfaces in a non sliding motion, while the viscous friction is dependent on properties of the oil layer formed in the sliding regime between the two contact surfaces.

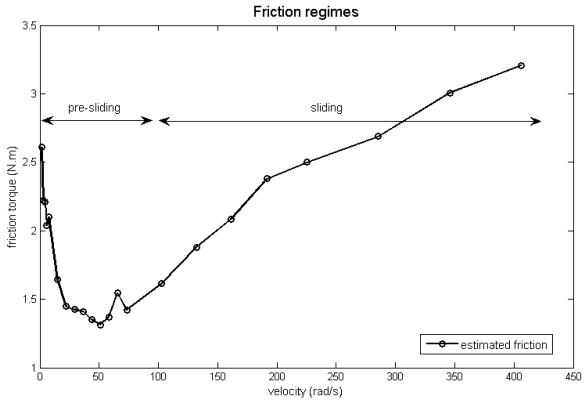


Figure 5.1: Friction regimes. Pre-sliding, friction phenomena that appears in the low velocities range; Sliding, friction phenomena in the greater velocities range.

5.2 Friction in robot joints

In robotics, the friction phenomena can be influenced by several variables and aspects, such as:

- Operational temperature
- Joint velocity
- Gearbox oil
- Robot joints configuration (different inertias)
- Motion positions (operational point)
- Load/tool coupled to the robot
- Usage
- Backlash (gear play)
- Robot model
- Different individuals for the same robot model

A successful fault detection method should generate alarms only when faulty states are present in the system, regardless disturbances. In this sense, it is vital when designing the fault detection method to have a great understanding of how the faults affect the parameter as well as the disturbances. For example, it is known that, before a gearbox breakdown, the friction phenomena increases significantly; also, that due to the use of lubricant oil, the temperature influence is relevant.

These informations will build up the basis to choose the estimated parameters to use with the change detection purpose. Nevertheless, it is a very challenging task to gather all this information. In general, before the decision of the design of a diagnosis system there is lack of information about the faults, its transiency and the effects in the system, as well as what should be called as the system normal operation behavior. Such knowledge is mainly acquired from long term experiments or domain experts.

The next sections of this chapter presents the result of an effort to gather knowledge about the friction behavior in robot joints through several of the mentioned variables.

5.3 Fault-free friction behavior in robot joints

Using the method described in Section 4.4.1 to estimate the friction curve, the aspects below have been studied in detail. They are here considered as to be fault-free since they are not directly related to joint malfunctions but to situations present in a normal operation of any industrial robot.

- *Operational point dependency*: due to asymmetries on the joints it is known that the friction phenomena is operational point dependent.

- *Presence of load*: the load presence obviously affects the robot inertial moments and therefore it is important to analyze its influence.

- *Robot joints configuration*: also affects the robot inertial moments but in a different way. If for example the new center of mass is too far from the moving axis, the link flexibilities may be more evident. Changes on configuration are also important because they can simulate how different tools attached to the robot would behave.

- *Oil*: there are different types of lubricant used in gearboxes, with different properties that are known to affect the friction phenomena, specially in the sliding regime.

- *Temperature Influence*: friction, specially the viscous part, is affected by the temperature operation of the robot.

5.3.1 Operational point, load, joints configuration

The friction behavior for the axis 2 of an ABB IRB 6650 industrial robot has been analyzed for changes in operational point, attached load and joints configuration.

Experimentation

A description of the experiments and its results can be found respectively in Figures 5.2 and 5.3. Each case has been realized with and without a load of $175kg$ coupled to the robot and in all experiments the robot was in thermal equilibrium with the environment temperature.

Concluding remarks

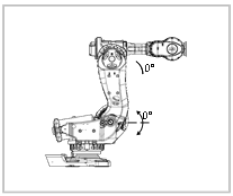
The analysis of the results of these experiments reveals the following about the friction phenomena:

- Operational point dependency: the friction phenomena did not change significantly with the different operational points.

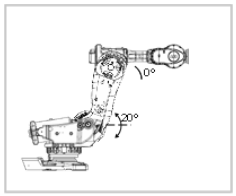
- *Robot joint configuration*: the friction phenomena did not change significantly with the different robot configurations.

- *Presence of load*: the presence of load affected the friction phenomena with more relevance for the lower velocities than for the higher ones.

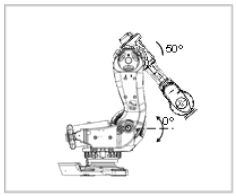
- *Gravity influence*: there is an increase of the friction phenomena in the cases where the moving joint is under a higher influence of the gravitational torques. This is easily seem when comparing the Operational Point A 5.2(b) (higher gravitational torques) and Configuration B 5.2(d) (lower gravitational torques).



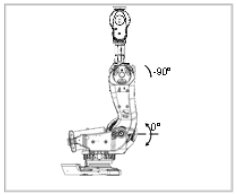
(a) Default. Movements around 0°; all other axes configurations in 0°.



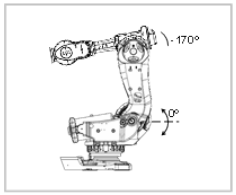
(b) Operational Point A. Movements around 20°; all other axes configurations in 0°.



(c) Configuration A. Movements around 0°; axis 3 configuration: 50°.

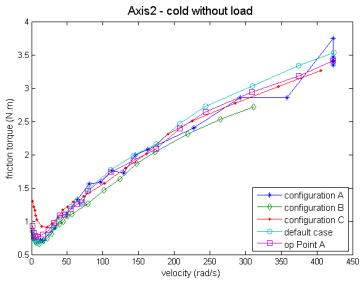


(d) Configuration B. Movements around 0°; axis 3 configuration: −90°.

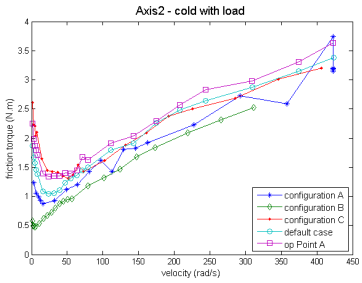


(e) Configuration C. Movements around 0°; axis 3 configuration: −170°.

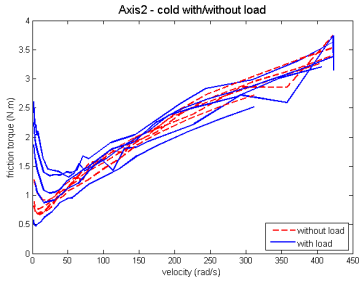
Figure 5.2: Experiment cases for axis 2.



(a) Experiments without load.



(b) Experiments with load.



(c) Comparison withwithout load.

Figure 5.3: Result of experiments for axis 2.

5.3.2 Oil and temperature

The friction phenomena in the sliding regime is dependent on properties of the oil layer formed between the contact surfaces. The property of an oil that may cause the biggest change in friction is its viscosity, it is natural to think that the friction will increase with the oil viscosity and, since the oil viscosity is sensible to temperature, that the oil temperature will also affect the friction phenomena.

In this work three oils with different viscosity and properties are considered:

- Oil 1
  - Kinematic viscosity at 40°C (mm^2/s): 150
  - Kinematic viscosity at 100°C (mm^2/s): 18
- Oil 2
  - Kinematic viscosity at 40°C (mm^2/s): 220
  - Kinematic viscosity at 100°C (mm^2/s): 34.4
- Oil 3
  - Kinematic viscosity at 40°C (mm^2/s): 320
  - Kinematic viscosity at 100°C (mm^2/s): 52.7

Note the different values of kinematic viscosity of the oils and its great sensibility to temperature.

Experimentation

The friction is monitored in an ABB IRB 6650 industrial robot. Since the temperature directly affects the viscosity of the oil, the friction curves are estimated with several different temperatures (varying from $24 - 50°C$).

The experiments are realized in the following routine:

- Warm-up cycle: the gearbox is warmed-up moving the axis from -50° to 50° with maximum speed for 30 min.
- Cooling down with friction estimation: after the gearbox reaches a certain temperature, the experiment to estimate the friction is realized several times with a 30 min interval between each other in order to have as many different temperatures as possible.

Environment and oil temperature measurements are available during the whole routine, which is repeated for the three different oils. Figure 5.4 shows the temperature measurements during one of the experiments with the warm-up and cooling down cycles.

Figure 5.5 shows the friction torque for each oil at the same gearbox temperature of $22°C$. Note that the *main changes appears in the sliding regime.*

For the same oil, the temperature influence is shown in Figure 5.6 where the gearbox temperature was varied from $24 - 50°C$. It is easy to realize the effects in friction caused by the different temperatures.

For comparison reasons, the friction torques for all three oils in the whole temperature range are plotted as surfaces in Figure 5.7.

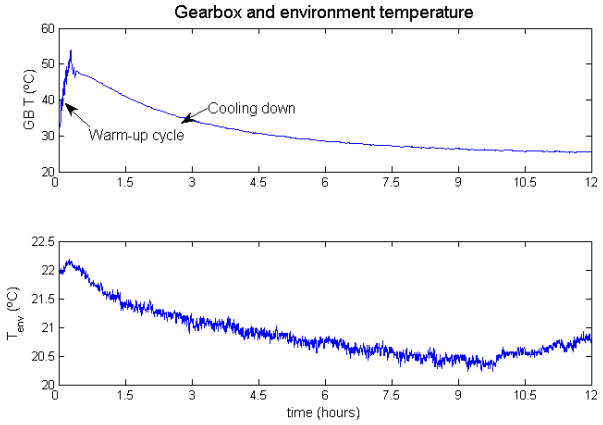


Figure 5.4: Gearbox and environment temperature measurements during the experiments.

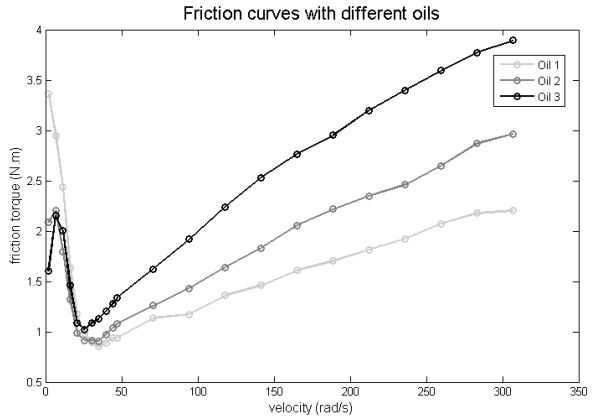


Figure 5.5: Estimated friction torques with different oils and same temperature.

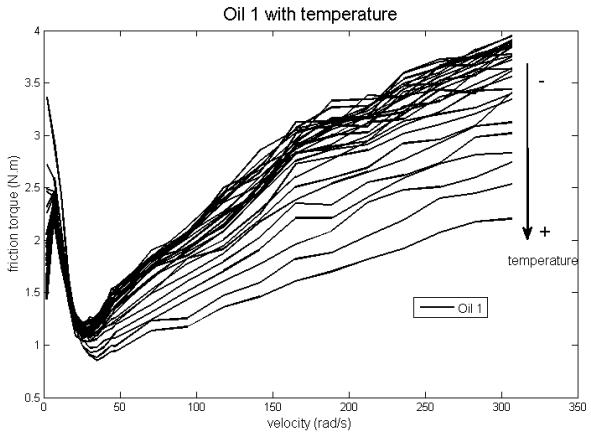


Figure 5.6: Estimated friction torques for same oil with different temperatures.

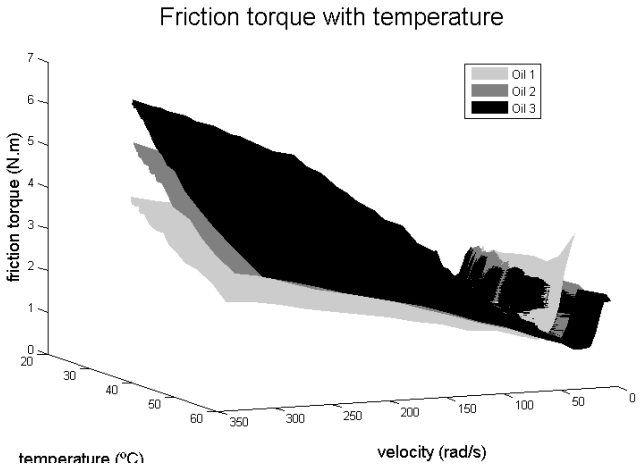


Figure 5.7: All surfaces plotted together.

Concluding remarks

The main effects caused by the use of different oils and temperature ranges are directly related to the viscosity of the oil, greater viscosities will cause an increase of the friction torques. Summarizing:

- *Oil*: greater friction torques with greater viscosities.
- *Temperature*: smaller friction torques with greater temperatures.
- *Shape of the curves*: the changes are mostly observed in the sliding regime. It is known that in this range the friction phenomena is highly dependent on the lubricant properties.

See Appendix A for more on temperature influence.

5.4 Faulty friction behavior in robot joints

In the context of fault detection, besides the knowledge of the normal variation of the observed phenomena, it is also important to understand how the fault affect the observations.

To gather the knowledge of how the faults affect the system, the friction was estimated in several worn-out robots. These robots had been operating for several years, starting to reach the end of their life time.

5.4.1 Wear

One of the common phenomena that appears before the breakdown of a gearbox is an increase of the friction. The phenomena is directly related to the fact that the gearbox metals start to peel with usage, increasing the free particles inside it and concomitantly the friction. At this point, the vibrations in the joint also increases.

The phenomena is shown in Figure 5.8 where it is obvious to identify the ill conditioned robot.

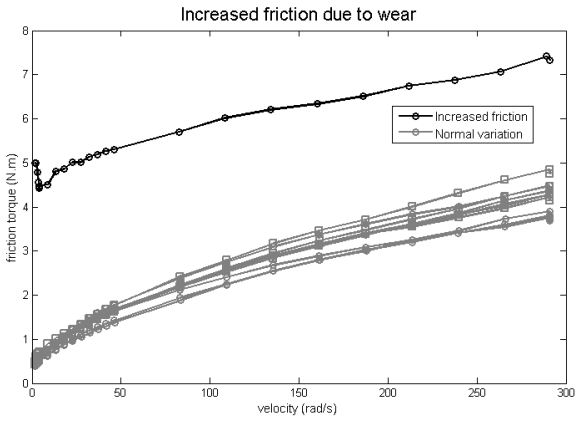


Figure 5.8: Increased friction phenomena caused by wear.

Transient behavior

Is is also important to understand how the fault develops with time. For this, extensive experiments were held in a robot which has been ran with a dedicated cycle to simulate the wear down process while the friction estimation experiment was realized concomitantly. The transient behavior of the fault caused by wear is shown in Figure 5.9;

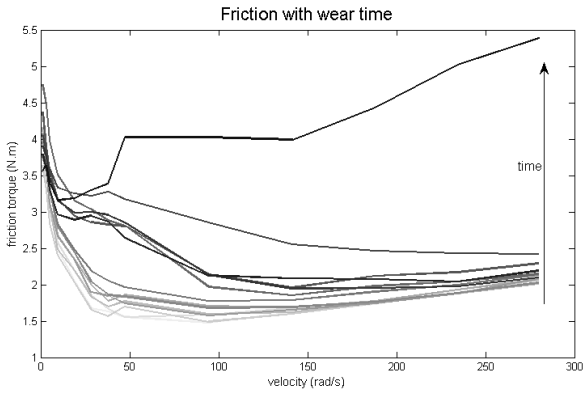


Figure 5.9: Transient behavior of the increased friction phenomena caused by wear.

5.4.2 Conclusions

It can be seem from Figure 5.9 that the influences appears with greater relevance in the pre-sliding regime, in this case for velocities lower than $100 rad/s$, and, of course, as an increase of the friction.

5.5 Concluding remarks

The results presented in this chapter are vital to define which parameters should be considered for the fault detection task. The observations from the faulty and fault-free system can be summarized as follows:

- Fault-free behavior
  - Inertial torques: the variables operational point, robot joint configuration and load presence affect the friction phenomena in a similar manner. For cases where there is an increase of the inertial torques to overcome either caused by a load, increased gravitational torque or arm inertia change, it is observed an increase of the friction phenomena, which is explained by the increase of the interactions forces between contact surfaces in the gearbox.
  - Gearbox lubricant: the use of more viscous lubricants will increase the friction phenomena, with greater relevance in the sliding regime.

  - Temperature: temperature was shown to be a very important variable in friction phenomena, affecting considerably the slopes in the sliding regime.

- Fault effects

  - Wear: the presence of free metals inside the gearbox caused by wear increases the friction phenomena.
  - Transient behavior: the friction increase with the wear down process, the phenomena has special relevance in the pre-sliding regime.

- Common observed phenomena

  - Shape of the curves: for the cases presented here, the shape of the friction curve did not vary considerably. Composed by a pre-sliding regime with high nonlinearities and the sliding regime where the curve approximates a line.
  - Individuality: even though not extensively presented in this chapter, it is important to note that the friction curves values varied with robot type, axis and robot individuals. This is important when delimiting the change detection range (for example, whether the same method and values can be used for all robot types or only for the same individual).

Considering the aspects presented until now, Chapter 6 presents a method to detect faults using friction as a parameter for fault detection.

Chapter 6

Friction change detection

The past Chapters have built the needed knowledge basis to perform the detection; Chapter 5 presented how the faults and disturbances affects the monitored friction phenomena while Chapters 3 and 4 presented methods and experiments to identify relevant parameters from the friction phenomena.

Using this knowledge, the objective of this chapter is to define a change detection method and to explore its performance over some case studies.

6.1 Definitions

Some definitions are reviewed in order to clarify the terms used.

6.1.1 Change detection

Given an estimation of a parameter $y_t = \theta_t + e_t$ where θ_t is deterministic and e_t additive noise, the task of change detection here is defined as finding a change (abrupt or incipient) in θ_t, occurring at *change time* k.

Figure 6.1 illustrates the change detection scheme used in this work. The task of estimating θ_t from y_t is referred as *estimation*. A residual ε_t is said as the difference between θ_o, considered as the normal value of θ, and its actual value θ_t. A *distance measure* is referred as a metric of the changes in the residual. Finally, a stopping rule is a filter that takes a distance measure as input and generates an alarm when θ_t has exceeded a certain threshold.

Figure 6.1: Change detection scheme

Several distance measures can be used, for example, changes in the mean or in the variance (see Section 2.3 for more).

There are also several stopping rules available, a classical approach is to apply a test in the distance measure using a threshold as an estimate of the standard deviation at each time, such that $|s_t| < 3\sigma_t$, this is also called as the

3 sigma test. For more robustness of the test, the variance using the non faulty data be included as $|s_t| < 3\sqrt{{\sigma_t}^2 + {\sigma_{o,t}}^2}$.

For signals with too high variation or for incipient faults, the 3 sigma test approach can generate too many false and/or missed alarms. The CUSUM, defined below, is a test used to improve the robustness of the stopping rule.

The CUSUM test

The CUSUM test statistic, see [18], is formulated by the following algorithm:

Algorithm 7 *CUSUM test*

$$\begin{aligned} g_t &= g_{t-1} + s_t - v \\ g_t &= 0, \ and \ \widehat{k} = t \ if \ g_t < 0 \\ g_t &= 0, \ and \ t_a = t \ and \ alarm \ if \ g_t > thold > 0. \end{aligned} \tag{6.1}$$

The test statistics g_t sums up its input s_t, with the idea to give an alarm when the sum exceeds a threshold. The drift variable v is set to compensate for the variation of the parameter caused by noise and errors in the estimation, while the threshold choice is related to the trade off between false alarms and detection time. Note that the test as defined in Algorithm 7 is a one-sided test, if s_t can be negative another CUSUM must be run in parallel.

6.1.2 Fault isolation - the hypothesis test

Fault isolation is the task of determining the kind, location and occurrence time of a fault. There are several ways to indicate the kind of fault, a very useful is the use of hypothesis test and decision structure as described by Nyberg in [20]. In this approach, several test quantities T_i (any quantity sensible to a fault) are estimated and compared in a hypothesis test to determine which kind of fault occurred.

A hypothesis test can be defined as the decision between the two states possible for a fault (present or not). A hypothesis test can be represented as a matrix where the rows are the hypotheses and the columns the test quantities, the elements of such matrix are the result of the test quantity for each hypothesis (each monitored behavioral mode of the system), a 0 value means that the test quantity does not relate to the hypothesis and an X that it can relate.

	T_0	T_1	T_2
F1	X	0	X
F2	X	X	0
NF	0	0	0

The illustration presents an example where T_0 can affect F1 and F2 and T_1 F2 and t_2 F1. NF stands for the non-fault hypothesis, where none of the test quantities affect it.

6.2 Building the detection scheme

In this Section each block/function in Figure 6.1 is defined for the detection method used.

6.2.1 Estimation

Four parameters are used to perform the detection.

- f_c: estimated Coulomb friction from the method defined in Section 4.4.3.

  f_c can be considered as a good measure of changes appearing in the pre-sliding regime of the friction curves.

- f_v: estimated Coulomb friction from the method defined in Section 4.4.3.

  f_v can be considered as a good measure of changes appearing in the sliding regime of the friction curves.

- $\int$: trapezoidal integral approximation of the friction characteristic curve.

  $\int$ can be considered as a good measure for detection of general changes in the friction curves, it also gives an idea of the amount of energy needed to overcome friction in the velocities range.

- $\int_{low}$: trapezoidal integral approximation of the friction characteristic curve for low velocities ($vel < vel_x$).

  $\int_{low}$ complements the f_c parameter in the sense that it carries the behavior of the friction curve for low velocities.

Defining the residual - moving average

A great challenge when defining a detection method is to choose the normal values of the monitored parameter, θ_0, having a good estimation of the real θ_0 usually requires long-term evaluation which is not always feasible. To deal with this, one could for example suppose that the estimated parameter inside a window of the whole data is said to be normal, $\widehat{\theta_0}(t-k1 : t-k2)$, while the last sample could be considered as the actual parameter, θ_t, this technique is also called as *sliding window* and is very important in fault detection.

Another approach, the one used here, is to take the estimates with a *moving average*, defined below.

$$\theta = \lambda y_{t-1} + (1-\lambda) y_t \tag{6.2}$$

For $\lambda < 0.5$, the result of the average gives more relevance for the last samples (the present estimations are more important). For $\lambda > 0.5$, the result of the average gives more relevance for the past samples (the past estimations are more important).

Considering that during the first samples the parameter is to be considered as normal, the fault free parameter, $\widehat{\theta_0}$ is averaged with $\lambda_0 > 0.5$. To improve the noise sensibility, the estimation of the actual parameter, $\widehat{\theta_t}$, are taken with $\lambda_t = 1 - \lambda_0$. In this manner, both signals are filtered, attenuating the noise influence while $\widehat{\theta_0}$ gives more relevance to past data and $\widehat{\theta_t}$ to the present data.

Let y_t be the estimated parameter, the residual is then taken as:

$$\begin{aligned} \lambda_0 &> 0.5 \quad \lambda_t = 1 - \lambda_0 \\ \widehat{\theta_0} &= \lambda_0 y_{t-1} + (1-\lambda_0) y_t \\ \widehat{\theta_t} &= \lambda_t y_{t-1} + (1-\lambda_t) y_t \\ \varepsilon_t &= \widehat{\theta_t} - \widehat{\theta_0} \end{aligned} \tag{6.3}$$

The *distance measure* is taken as the residual from Equation (6.3), $s_t = \varepsilon_t$.

6.2.2 Stopping rule

The CUSUM test is used as a stopping rule. The drift constant is set as $v_t = k_v \widehat{\theta_t}$ ($k_v > 0$), in this manner the test is reduced to:

$$\begin{aligned} g_t &= g_{t-1} + \varepsilon_t - v \\ g_t &= g_{t-1} + (\widehat{\theta_t} - \widehat{\theta_0}) - k_v \widehat{\theta_0} \\ g_t &= g_{t-1} + \widehat{\theta_0} - (1 + k_v)\widehat{\theta_0} \end{aligned} \tag{6.4}$$

Configuring the test in this manner, k_v becomes a measure of the smallest change to be detected. For example, supposing no error in the estimation of $\widehat{\theta_0}$ and $\widehat{\theta_t}$, the test g_t will only increase if $\widehat{\theta_t}$ is at least $k_v\widehat{\theta_0}$ greater than $\widehat{\theta_0}$, in other words, the parameter has to deviate k_v from normality.

However, since there are errors in the estimation of $\widehat{\theta_0}$ and $\widehat{\theta_t}$, it can also be included the normal variation of the parameters in the drift as $v = k_v\widehat{\theta_0} + \alpha$. A good choice for α is to take a scaled value of the standard deviation of the residual, $\alpha_t = k_\alpha \sigma(\varepsilon_{1:t})$.

The threshold can now be defined as $thold_t = 3v_t$.

6.2.3 Choosing the hypothesis

There are three behavioral modes suitable to be separated using friction estimated parameters:

NF - no fault : normal situation.

H0 - increased friction : the friction increase as an indication of a possible near breakdown.

H1 - high increased friction : the friction increase when the gearbox is close to a breakdown. This situation can be considered as an urgent case and maintenance should be proceeded immediately.

The most important to detect is H1, which can compromise the whole system where the robot is operating. H0 can be seem as a pre-alarm for H1.

The test quantities used to relate these behavioral modes are:

T1 - increase of friction : the parameter $\int_{low}$ is used to give an indication of an increased friction, the indication is true if the last two samples exceeded the threshold.

$$\int_{low}(t-1:t) > thold$$

T2 - high increase of friction : the parameter $\int$ has to exceed the threshold for at least the three last samples. Since the main observed changes occur in the low velocities range it also checks $\int_{low}$ and f_c.

$$\int(t-2:t) \quad \text{AND} \quad \int_{low}(t) \quad \text{AND} \quad f_c(t) > thold$$

T0 - no alarm : all other cases.

The matrix representing the hypothesis test becomes:

	T0	T1	T2
H0	0	X	0
H1	0	0	X
NF	X	0	0

6.3 Case studies

In this section the detection is evaluated within some case studies. The cases are generated with real data and/or simulations using real data. The same detection algorithm was used for all cases with $\lambda = 0.8$, $k_v = 0.01$ and $k_\alpha = 0.4$.

6.3.1 Case 1: Normal operation

Objective

Check the behavior of the detection method in the case where there is no fault in the system.

Experiment

The friction experiment was realized several times in an industrial robot which was considered to be operating normally. The conditions of temperature and motions (trajectories for the identification) were kept constant during the experiments. Figure 6.2 shows the results of the detection for the case.

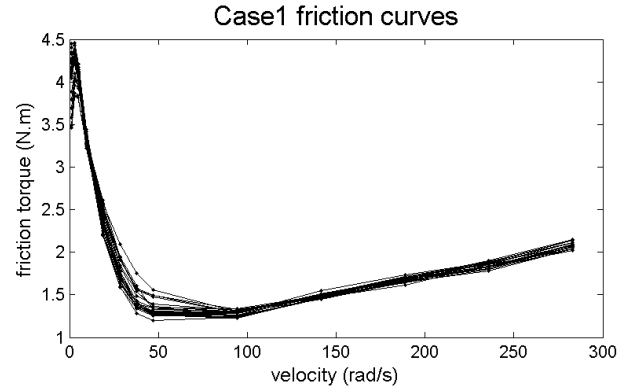
Conclusions

It can be seem that even though f_c and f_v exceeded the threshold, the use of the hypothesis test assited to consider the data as normal. To improve the performance of these parameters, v should be increased tuning k_v andor k_α.

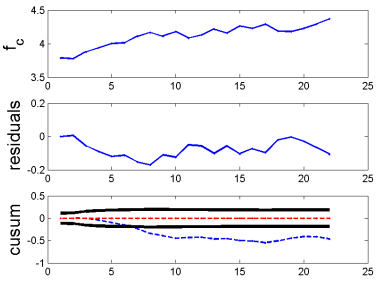
6.3.2 Case 2: Gearbox breakdown

Objective

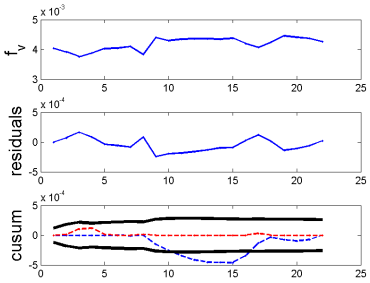
Check the behavior of the detection method in a case with a real breakdown process in the gearbox.



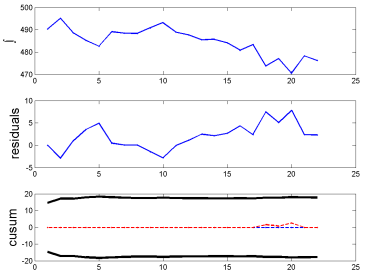
(a) Friction curves



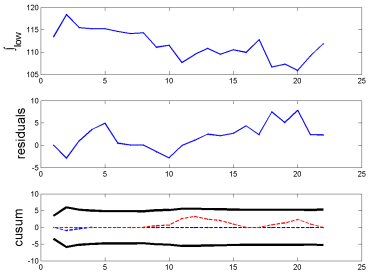
(b) f_c



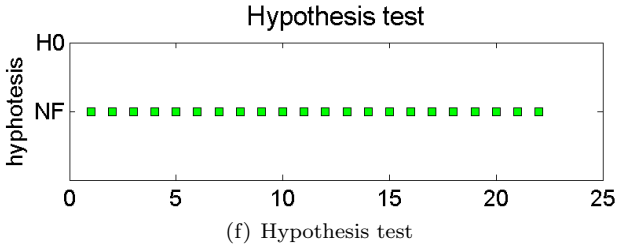
(c) f_v



(d) $\int$



(e) $\int_{low}$



(f) Hypothesis test

Figure 6.2: Case 1: normal variation.

Experiment

The data used in the detection is the same as the one used to generate Figure 5.9 in which a gearbox has been ran continuously to force a wear down process while the friction estimation experiment was realized, temperature and motions were kept constant during the experiments. The results are shown in Figure 6.3.

Conclusions

In this case, both hypothesis H0 (increase of friction) and H1 (high increase of friction) are excited. H0 is excited one sample before H1, giving an indication of a possible breakdown, while H1 is excited when the gearbox is close collapse (the gearbox broke after the last sample used in this data set).

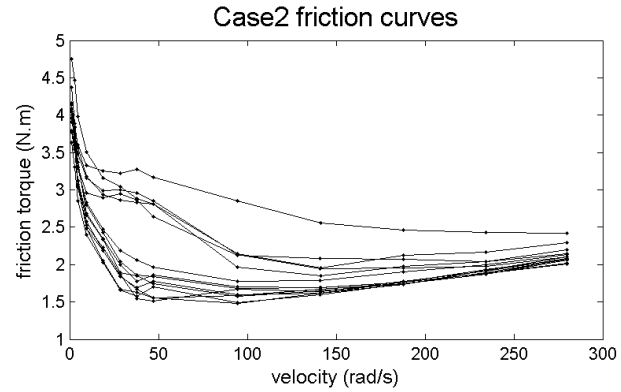
6.4 Concluding remarks

In this chapter a fault detection scheme has been presented and evaluated within some scenarios. The detection algorithm proposed uses several tools to improve the robustness of the test and reduce the number of parameters to be tuned to three:

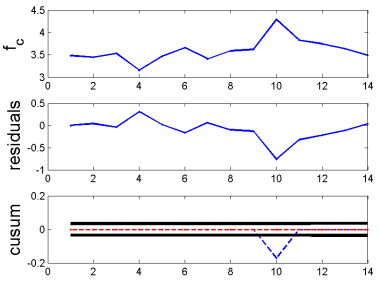
- λ: compromise between fault size, detection time and filter attenuation.
- k_α: compromise between robustness of the test and detection time.
- k_v: percentage of the size of the fault to be detected.

The tuning of these parameters are quite straightforward and does not require too many attempts. The use of the hypothesis test is also very important, improving the robustness of the test and providing fault isolation.

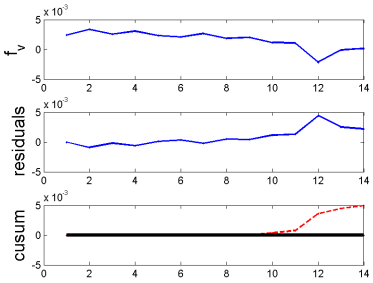
Finally, the scenarios tried to approach reality as possible to validate the approach, two cases were presented here, but the method was tested in several others and the detection performed well in all of them.



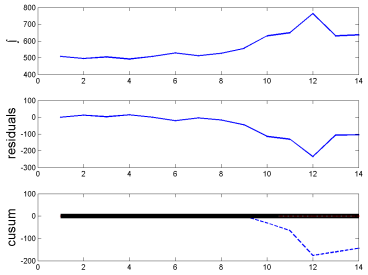
(a) Friction curves



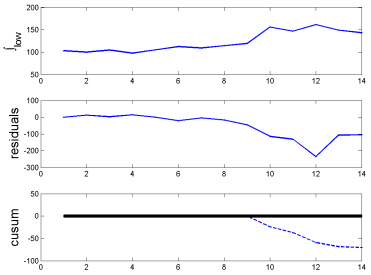
(b) f_c



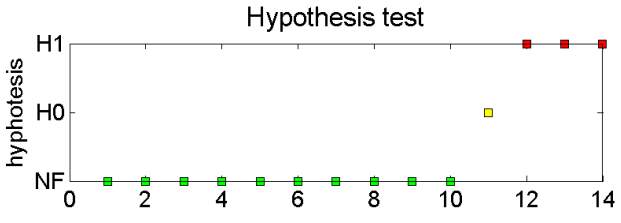
(c) f_v



(d) $\int$



(e) $\int_{low}$



(f) Hypothesis test

Figure 6.3: Case 2: gearbox breakdown.

Chapter 7

Conclusions

This Chapter summarizes the thesis report and leave comments on future work.

7.1 Summary

A review on fault detection was first presented introducing the field to the reader and discussing the unavoidable trade offs present in any detection scheme. It also pointed out the advantages and disadvantages between the several methods available for the application sought, friction change detection. It has been shown that for detection of multiplicative faults (as in the case of friction), the detection using parameter estimation is more convenient than state estimation.

Chapter 3 introduced the Robotics field, discussing the main phenomena and restrictions. It presented some robot and friction models suitable for a parameter based detection. Chapter 4 complemented it with identification methods and experiments of an industrial robot arm. The characteristics of the model and identification methods were also considered for the application sought, assisting the choice of the estimation method.

One of the most important contributions of the thesis was presented in Chapter 5, the analysis of friction behavior under several variables. It has been shown that the most important disturbance on the friction is caused by the viscosity of the oil used inside the gearbox, which varies significantly with the operational temperature. The experiments realized with faulty robots revealed the behavior of the friction phenomena under a wear down process which is an important phenomena to the operation of a joint. The knowledge gathered from this chapter is not only vital for the diagnosis scheme proposed but also relevant for other tasks such as control, modeling and simulations.

With the knowledge basis built with the other chapters, Chapter 6 finally presented and defined a fault detection method based on estimated friction parameters. The test robustness was achieved with the use of several tools such as the moving average in the estimation and the set of the drift variable in the CUSUM test. Also, fault isolation was achieved with the use of a hypothesis test. The method was evaluated in several scenarios (two were presented here) generated from real data, showing its validity.

Finally, Appendix A, presented a more detailed analysis of the friction behavior under the disturbance of temperature, providing insights for improving

the robustness of the experimentation.

7.2 Future work

There are several possible improvements on the work presented, some of them are listed below:

- Define the requirements of the method: even though the detection performed well in the scenarios presented, some further investigation on the experimentation rate and conditions (temperature and oil for instance) should be considered to provide a robust tool suitable to be used in any case (different robot models with different operational conditions).
- Extensive evaluation of the method: it would be interesting to check the performance of the method in many cases as possible, specially in field applications, to testify its validity as a tool suitable for condition monitoring and maintenance assistance.
- Automated detection tuning: the detection method presented in Chapter 6 needs at least three parameters to be tuned, and even though tuning them is not a complex task, it is more suitable for the user if no tuning is needed. This could be achieved, for example, by checking the variance of the parameters for normal robots and building a database with tuning parameters for every robot type.
- Estimate the detection time: not covered in the work is the task of estimation the time that a fault appeared in the system. This feature is important to keep track of the changes and therefore assist maintenance. There are several methods used to estimate the change time, found for example in [18].
- Estimation of the remaining life time of a robot: an interesting application of the detection is to try to estimate the remaining life time of a robot. This could be achieved for example by checking the size of the faults and relating them to a life time model.
- On-line parameter estimation: the method proposed here to estimate the friction parameters requires a dedicated experiment, which means that in the field, the robot needs to be stopped to run the test cycles. This situation is not desirable since the robots usually play an important role in their application. In this sense, it would be interesting to define a friction estimation method using data from a normal cycle of the robot (logged from a usual workcyle). However, this is a very challenging task since the parameters can only be identified if their frequency range is excited, which could not be the case with a general input.

Bibliography

[1] H. Olsson, K.J. Åström, C. Canudas de Wit, M. Gäfvert and P. Lischinsky : *Friction Models and Friction Compensation*, European Journal of Control, **4**, (1998), 176-195.

[2] B. Feeny and F. Moon: *Chaos in a forced dry-friction oscillator: Experiments and numerical modeling*, Journal of Sound and Vibration, **170**3, (1994), 303-323.

[3] R. Isermann: *Supervision,fault-detection and fault-diagnosis methods - An Introduction*, Control Engineering Practice, **5**5, (1997), 639-652.

[4] R. Isermann and P. Ballé : *Trends in the application of model-based fault detection and diagnosis of technical processes*, Control Engineering Practice, **5**5, (1997).

[5] R. Isermann and B. Freyermuth: *Process fault diagnosis based on process model knowledge - Part I: Principles for fault diagnosis with parameter estimation*, ASME J. of Dynamic Systems, Measurement, and Control, **113**, (1991a), 620-626.

[6] R. Isermann and B. Freyermuth: *Process fault diagnosis based on process model knowledge - Part II: Case study, Experiments*, ASME J. of Dynamic Systems, Measurement, and Control, **113**, (1991a), 620-626.

[7] R. Isermann: *Process Fault Detection Based on Modeling and Estimation Methods-A Survey*, Automatica, **20**4, (1984), 387-404.

[8] R. Isermann: *Estimation of physical parameters for dynamic processes with application to an insdustrial robot*, Int J. Control, **55**, (1992), 1287-1298.

[9] R. Isermann: *Fault Diagnosis of machines via parameter estimation and knowledge processing - Tutorial paper*, Automatica, **29**4, (1993), 815-835.

[10] Rabinowicz, E.: *The nature of static and kinetic coefficients of friction*, Journal of Applied Physics, **22**11, (1951), 1373-79.

[11] L. Márton and B. Lantos: *Modeling, Identification, and compensation of stick-slip friction*, IEEE Transactions on Industrial Electronics, **54**1, (2007), 511-21.

[12] B. Bona, M. Indri and N. Smaldone: *Nonlinear friction estimation for digital control of direct-drive manipulators*, **22**11, (1951), 1373-79.

[13] P. B. Hvass and D. Tesar: *Condition Based Maintenance for intelligent electromechanical actuators*, Robotics Research Group of the University of Austin at Texas, (2004).

[14] F. Caccavale: *Experiments of Oberver-based Fault Detection for an Industrial Robot*, Proceedings of the 1998 IEEE International Conference on Control Applications, (1998), 480-4 vol.1.

[15] F. Caccavale and ID. Walker: *Oberver-based Fault Detection for Robot Manipulators*, Proceedings. 1997 IEEE International Conference on Robotics and Automation, (1997), 2881-7 vol.4.

[16] H. Schneider and P. M. Frank: *Oberver-based supervision and fault detection in robots using nonlinear and fuzzy logic residual evaluation*, IEEE Transactions on Control Systems Technology, **4**3, (1996), 274-82.

[17] M. Saif: *Robust discrete time observer with application to fault diagnosis*, IEEE Proceedings Control Theory and Applications, **145**3, (1998), 353-7.

[18] F. Gustafsson: *Adaptative filtering and change detection*, John Wiley & Sons Ltd., (2000).

[19] L. Ljung and T. Söderström: *Theory and practice of recursive identification*, The MIT Press, (1983).

[20] M. Nyberg: *MModel Based Fault Diagonsis: Methods, Theory, and Automotive Engine Applications*, PhD thesis, Linköpings universitet, Linköping, Sweden. Linköping Studies in Science and Technology, (1999). Dissertation No. 591.

[21] P. C. Young: *Parameter estimation for continuous-time models - a survey*, Automatica, **17**23, (1981).

[22] V. F. Filaretov, M. K. Vukobratovic and A. N. Zhirabok: *Parity relation approach to fault diagnosis in manipulation robots*, Mechatronics, **13**2, (2003), 141-52.

[23] Q. Wang, Q. Bi and B. Zou: *Parameter identification of continuous-time mechanical systems without sensing accelerations*, Computers in Industry, **82**, (1996), 207-217.

[24] E. Wernholt: *On Multivariable and Nonlinear Identification of Industrial Robots*, Licentiate thesis, Linköpings universitet, Linköping, Sweden. Linköping Studies in Science and Technology, (2004), Thesis No. 1131.

[25] E. Wernholt: *Multivariable Frequency-Domain Identification of Industrial Robots*, PhD thesis, Linköpings universitet, Linköping, Sweden. Linköping Studies in Science and Technology, (2007), Thesis No. 1138.

[26] M. Östring: *Identification, diagnosis, and control of a flexible robot arm*, Licentiate thesis, Linköpings universitet, Linköping, Sweden. Linköping Studies in Science and Technology, (2002). Thesis No. 948.

[27] L.-W. Tsai: *Robot Analysis - The Mechanics of Serial and Parallel Manipulators*, John Wiley & Sons Ltd., (1999).

[28] M. Nordin and P.-O. Gutman: *Controlling mechanical systems with backlash - a survey*, Automatica, **38**, (2002), 1633-49.

[29] G. Hovland, S. Hanssen, E. Gallestey, S. Moberg, T. Brogårdh, S. Gunnarsson and M. Isaksson: *Nonlinear Identification of backlash in robot transmissions*, ABB Corporate Research.

[30] J. Holtz and L. Springob: *Identification and compensation of torque ripple in high-precision permanent magnet motor drives*, IEEE Trans. on Industrial Electronics, **43**2, (1996), 309-320.

[31] L. Ljung: *System Identification - Theory for the user*. Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd Edition, (1999)

[32] M. Norrlöf: *Modeling of industrial robots*, Technical Report LiTh-IsY-R-2208, Department of Electrical Engineering, Linköping University, (1999).

[33] E. Andersson: *Real time thermal model for servomotor applications*, Technical Report, ABB Corporate Research, Sweden. (2006).

[34] J. Gunnar: *Dynamical analysis and system identification of the Gantry-Tau parallel manipulator.*, Master Thesis Report LiTh-IsY-EX–05/3762, Department of Electrical Engineering, Linköping University, (2005)

Appendix A

More on temperature influence

As seem in Chapter 5, the temperature affects considerably the friction phenomena in the sliding regime. A measure of how the friction is affected by the temperature is the integral of the friction curve, since the increase of temperature reduces the friction torques in the sliding regime, the integral will also decrease.

Figure A.1 shows the integral of the friction curves (computed using a trapezoidal approximation) for each oil as a function of temperature.

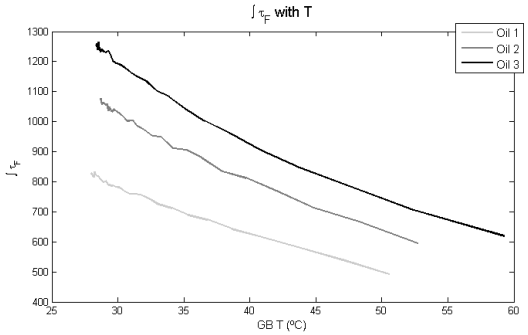


Figure A.1: $\int \tau_F$ with temperature.

From Figure A.1 it is easy to see that the changes present a *linear behavior with temperature* and with a similar rate. The estimated inclination of each curve in Figure A.1 are:

- Oil 1: -15.09
- Oil 2: -20.53
- Oil 3: -22.89

This characteristic indicates that *the oil properties that affect friction vary with similar rate and quite linearly with temperature* for all oils tested in this work (which are the ones recommended by the manufacturer).

A.1 Cooling down curve

If the cooling down curve for each oil is plotted together with the same initial temperature, as in Figure A.2, it is easy to realize that it is not dependent on the oil used inside the gearbox (the oils may have same thermal capacity), but only on the joint properties.

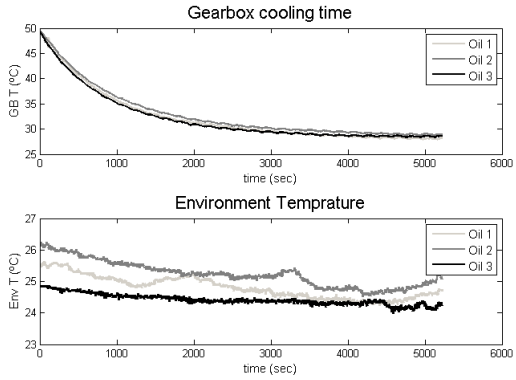


Figure A.2: Cooling time inside the gearbox and environment temperature.

A simple model of the temperature decay inside the gearbox can be derived, which is also known as the Newton law of cooling, described by Equation (A.1).

$$\hat{T(t)} = T_{env} + (T_o - T_{env})e^{-rt} \tag{A.1}$$

Where $\hat{T(t)}$ is the estimated temperature inside the gearbox, T_{env} is the environment temperature, T_o is the initial temperature of the gearbox and r is a constant characteristic of the system. Equation (A.1) only includes the conduction heat transfer and relies on the simplification that the body has relatively high internal conductivity, such that (to good approximation) the entire body is at same uniform temperature as it is cooled from the outside by the environment.

The constant r from Equation (A.1) was estimated for each oil as minizing the prediction error of the model, $e = \hat{T(t)} - T$. The estimated constant r for each oil is:

- Oil 1: $6.3664e-005$
- Oil 2: $5.9361e-005$
- Oil 3: $6.3340e-005$

The similar values of r indicate that the use of a global model of the cooling down curve is possible; the model approximation taking the mean of r for each oil ($r = 6.2122e-005$) is shown in Figure A.3, where is easy to see that this can be considered as a good approximation.

The fact that the gearbox temperature does not converges to the environment temperature indicates that there is a heat source inside the gearbox, in this case it is caused by an intermittent friction estimation cycle ran each 30

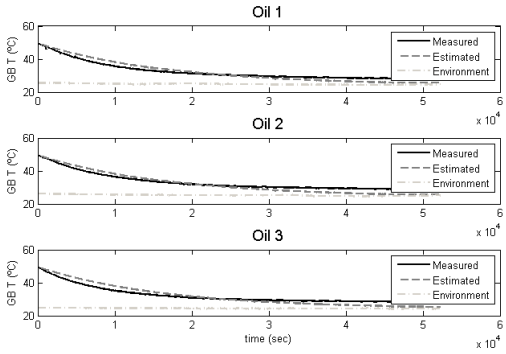


Figure A.3: Gearbox cooling down measurements and model.

min interval, keeping the gearbox temperature greater than the environment. Since this heat source is not included in the model, there is an error of $\approx 4°C$ from the model and measurements in steady-state.

A.2 Temperature increase with work cycle

Since the only heat sources in the system are the losses (friction for example) occurring in the robot, which are directly related to the work cycle $\mho$, the difference of temperature between the gearbox and the environment will be a function of $\mho$ and the rate that it is repeated. If the robot is run continuously with the same work cycle $\mho^*$, then the increase of temperature could be considered as:

$$T_o = T_{env} + T^*_{\mho} \tag{A.2}$$

Where $T^*_{\mho}$ is the added temperature to the system due to the losses in the work cycle $\mho^*$. In this manner, the only effect of T_{env} is an offset change in the cooling time curve.

Supposing that the robot always operates with the same work cycle so that $T_{\mho} = cte$ in steady-state and that the experiments are held after the same time since the robot was stopped, such that $T_{decay} = cte$, the gearbox temperature will only be function of T_{env}:

$$\begin{aligned} T_{GB}(T_{env}) &= T_o - \overline{T_{decay}} \\ T_{GB}(T_{env}) &= \overline{T_{\mho}} - \overline{T_{decay}} + T_{env} \end{aligned} \tag{A.3}$$

It is important to note, however, that this conclusion supposes that the losses in the robot remains constant with temperature, which is not true but could be used as a fair approximation. It is known for example, that there is a temperature dependency of the winding resistance for permanent magnet motors as:

$$R = R_{20°C}\frac{235 + T}{235 + 20} \tag{A.4}$$

A.3 Environment temperature influence

The environment temperature will of course cause differences in the estimation. The changes in the environment temperature can be caused by several factors, different stations in a year for example, and adds a need of considering the rate of the service routines for fault detection.

For example, supposing that the environment temperature vary $\pm 10\,^{\circ}\mathrm{C}$ during a year, it is expected the friction curve also to vary during the period. Therefore, depending on the periodicity of the experiments, the changes between the experiments will vary.

In Figure A.4 the rate of the experiments is simulated by taking different deltas of temperature for each set of experiments, a higher periodicity is simulated taking a delta of $2\,^{\circ}\mathrm{C}$ between two consecutive curves (black curves) and a lower rate with a $5\,^{\circ}\mathrm{C}$ difference (gray curves).

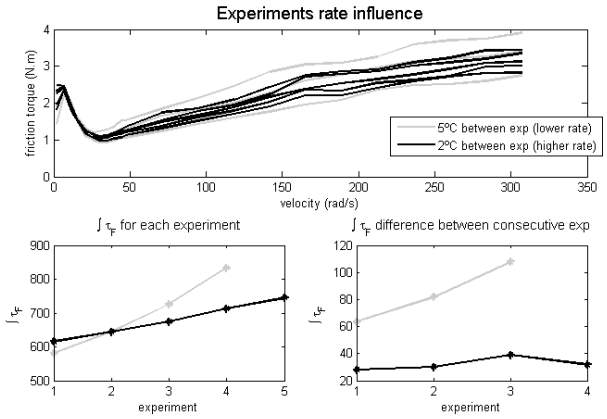


Figure A.4: Experiments rate influence. Note the bigger differences between two consecutive experiments whether the rate is lower.

The friction curve integral, $\int \tau_F$, is plotted in the same figure and the difference of the $\int \tau_F$ between consecutive experiments. It is easy to realize that the higher the rate of the experiments is, the lower the changes between the experiments will be, which is an important characteristic when tunning the detection algorithm.

A.4 Concluding remarks

From the analysis above, the following conclusions can be drawn about the temperature influence in friction:

- Cooling down model: the observed cooling down curves were practically the same for all oils tested.
- Work cycle influence: it was presented the suggestion that the increase of the gearbox temperature in the same environment temperature is due to the losses occurring in the work cycle, therefore, if the work cycle used

is the same, *the temperature of the robot would only be a function of the work cycle, the environment temperature and the decay time.*

- Experimentation rate: it has been shown that due to fluctuations in the environment temperature, *the rate of the experiments affect the changes appearing in friction between consecutive experiments.*