PERGAMON

Control Engineering Practice 9 (2001) 1361–1372

CONTROL ENGINEERING PRACTICE

www.elsevier.com/locate/conengprac

Extended Kalman filtering and weighted least squares dynamic identification of robot

M. Gautier[a,]*, Ph. Poignet[b]

[a] Institut de Recherche en Communication et Cybernétique de Nantes, IRCCyN, UMR, CNRS 6597, 1 rue de la Noe, BP 92101, 44321 Nantes Cedex 03, France

[b] Laboratoire d'Informatique, de Robotique et de Micro-électronique de Montpellier, LIRMM, UMR, CNRS 5506, 161 rue Ada, 34392 Montpellier Cédex 5, France

Received 2 February 2001; accepted 21 May 2001

Abstract

This paper presents an experimental comparison between the weighted least squares (WLS) estimation and the extended Kalman filtering (EKF) methods for robot dynamic identification. Comparative results and discussion are presented for a SCARA robot, depending on a priori knowledge and data filtering.

Keywords: Identification; Robot dynamics; Kalman filter; Least squares estimation

1. Introduction

In the class of estimation algorithms, the weighted least squares estimation (WLS) method (see Section 2) takes a particular place for robot manipulators identification. Based on the use of the inverse model linear in relation to the parameters, it allows to estimate the base inertial parameters providing measurement or estimation of the joint torques and the joint positions (Gautier, 1986; Gautier & Khalil, 1988; Gautier, 1990; Canudas de Wit, Siciliano, & Bastin, 1996; Kozlowski, 1998; Khalil & Dombre, 1999). Some alternative solutions for robust estimation in the class of least squares estimators have been proposed. Swevers, Ganseman, Tükel, Schutter, and Van Brussel (1997) formulated an approach based on the maximum-likelihood parameter estimation. But in the practical case, they are considered additive noise which lead to a WLS estimation. Califiore and Indri (1999) considered linear matrix inequalities (LMI) to take into account uncertainties in the observation matrix due to error modeling or measurement noise. An alternative method, more common in the automatic control community, is the use of Kalman filtering algorithm (see Section 4). Based on the direct dynamic model (the state space model) which is non-linear in relation to the state and the parameters, an extended state containing the physical parameters is considered. Both the state (position and velocity) and the parameters are estimated through only one model, considering the uncertainties in the model and the measurements. Guglielmi, Jonker, and Piasco (1987) applied the EKF approach to a SCARA robot, but without any comparison to LS. Gautier, Janin, and Presse (1993) have presented a comparison between the LS energetic and the EKF approaches, but it is based on the experimental identification of a single joint robot. This paper presents an experimental comparison on a two degrees of freedom robot, using WLS dynamic and EKF approaches. It is organized as follows: Section 2 introduces the dynamic model and the WLS estimation, Section 3 presents the extended Kalman filtering, Section 4 contains the description of the SCARA robot and Section 5 exhibits the comparisons of experimental results and discussion concerning the use of both algorithms.

2. Weighted least squares estimation

2.1. Identification model

The inverse dynamic model of a rigid robot composed of n moving links calculates the motor torque vector $\boldsymbol{\tau}$

*Corresponding author. Fax: +33-2-40-37-69-30.
E-mail addresses: maxime.gautier@irccyn.ec-nantes.fr (M. Gautier), poignet@lirmm.fr (P. Poignet).

PII: S0967-0661(01)00105-8

(the control input) as a function of the generalized coordinates (the state vector and its derivative). It can be obtained from the Lagrangian or Newton Euler equation as recalled here (Canudas et al., 1996; Khalil & Dombre, 1999; Kozlowski, 1998):

$$\boldsymbol{\tau} = \mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{N}(\mathbf{q}, \dot{\mathbf{q}}), \tag{1}$$

where $\boldsymbol{\tau}$ is the $(n\times 1)$ motor torque vector, $(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ the $(n\times 1)$ vectors of generalized joint positions, velocities and accelerations, respectively, $\mathbf{M}(\mathbf{q})$ the $(n \times n)$ inertia matrix, $\mathbf{N}(\mathbf{q}, \dot{\mathbf{q}})$ the $(n\times 1)$ vector of centrifugal, Coriolis, gravitational and friction torques.

Eq. (1) can be rewritten as a linear relation to a set of standard dynamic parameters $\boldsymbol{\chi}_S$ (Mayeda, Yoshida, & Osuka, 1990; Gautier & Khalil, 1990)

$$\boldsymbol{\tau} = \mathbf{D}_S(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})\boldsymbol{\chi}_S, \tag{2}$$

where $\boldsymbol{\chi}_S$ is the $(13n\times 1)$ vector of standard dynamic parameters:

$$\chi_S^j = [XX_j \; XY_j \; XZ_j \; YY_j \; YZ_j \; ZZ_j \; MX_j \; MY_j \; MZ_j \; M_j \; Ia_j \; F_{Vj} \; F_{Sj}]^T.$$

It is composed, for each link j, of $(XX_j, XY_j, XZ_j, YY_j, YZ_j, ZZ_j)$, the six components of the inertia tensor; (MX_j, MY_j, MZ_j), the three components of the first moment; M_j, the mass; I_{aj}, the total inertia moment for rotor actuator and gears; and F_{Vj}, F_{Sj}, the Coulomb and viscous friction parameters.

It has been shown that the set of standard dynamic parameters can be simplified to obtain the base inertial parameters. The base inertial parameters are defined as the minimum parameters which can be used to calculate the dynamic model. They represent the set of p parameters which can be identified using the dynamic model. These parameters can be obtained from the standard inertial parameters by eliminating those which have no effect on the dynamic model and by regrouping some others in linear relations. Symbolic and numerical solutions have been proposed for any open or closed chain manipulator (Mayeda et al., 1990; Gautier, 1990; Gautier & Khalil, 1990; Gautier, 1991; Khalil & Dombre, 1999) to obtain a minimal dynamic model

$$\boldsymbol{\tau} = \mathbf{D}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})\boldsymbol{\chi} = \sum_{i=1,\,\mathrm{p}} \mathbf{D}_{:,i}\chi_i. \tag{3}$$

The coefficients of the matrices $\mathbf{D}_S$ and $\mathbf{D}$ can be automatically calculated using a customized symbolic method (Khalil & Creusot, 1997; Khalil & Dombre, 1999).

2.2. Identification method

Usually, $\boldsymbol{\chi}$ is estimated as the least squares (LS) solution of an overdetermined $(r\times p)$ linear system obtained from sampling and filtering the dynamic model (3) along a trajectory $(\mathbf{q}(t), \dot{\mathbf{q}}(t), \ddot{\mathbf{q}}(t))$:

$$\mathbf{y}(\boldsymbol{\tau}) = \mathbf{W}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})\boldsymbol{\chi} + \boldsymbol{\rho}, \tag{4}$$

where $\mathbf{y}$ is the $(r\times 1)$ measurement vector, $\mathbf{W}$ the $(r\times p)$ observation matrix, and $\boldsymbol{\rho}$ the $(r\times 1)$ vector of errors.

It is considered to be a zero mean additive independent noise, with standard deviation σ_ρ such that

$$\mathbf{C}_{\boldsymbol{\rho\rho}} = E(\boldsymbol{\rho}^T\boldsymbol{\rho}) = \sigma_\rho^2\mathbf{I}_r, \tag{5}$$

where E is the expectation operator, $\mathbf{I}_r$ the $(r\times r)$ identity matrix.

In fact, $\mathbf{y}$ is obtained from the concatenation of n measurements vectors $\mathbf{y}^j$ of the n motor torques with different errors standard deviations.

Matrices $\mathbf{y}$ and $\mathbf{W}$ are sorted in order to regroup the rows of the joint j equation:

$$\tau_j = \mathbf{D}_{j,:}\boldsymbol{\chi},$$

where $\mathbf{D}_{j,:}$ is the row j of $\mathbf{D}$,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}^1 \\ \vdots \\ \mathbf{y}^n \end{bmatrix}, \quad \mathbf{y}^j = \begin{bmatrix} \tau_j(1) \\ \vdots \\ \tau_j(r/n) \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} \mathbf{W}^1 \\ \vdots \\ \mathbf{W}^n \end{bmatrix},$$

$$\mathbf{W}^j = \begin{bmatrix} \mathbf{D}_{j,:}(1) \\ \vdots \\ \mathbf{D}_{j,:}(r/n) \end{bmatrix}, \quad \boldsymbol{\rho} = \begin{bmatrix} \boldsymbol{\rho}^1 \\ \vdots \\ \boldsymbol{\rho}^n \end{bmatrix}.$$

An improvement in the ordinary LS solution is to calculate the WLS solution of the global system (4) (Gautier, 1997). The r^j rows corresponding to joint j equation are weighted by the coefficient of the error covariance diagonal matrix factorized as follows:

$$\mathbf{C}_{\boldsymbol{\rho\rho}} = (\mathbf{G}^T\mathbf{G})^{-1}, \quad \mathbf{G} = \mathbf{diag}(\mathbf{s}), \tag{6}$$

where $\mathbf{G}$ is a $(r\times r)$ diagonal matrix with the elements of $\mathbf{s}$ on its diagonal,

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}^1 & \cdots & \mathbf{s}^n \end{bmatrix}, \quad \mathbf{s}^j = \begin{bmatrix} \dfrac{1}{\hat{\sigma}_p^j} & \cdots & \dfrac{1}{\hat{\sigma}_p^j} \end{bmatrix},$$

where $\mathbf{s}^j$ is a $(1\times r^j)$ row matrix.

An unbiased a posteriori estimation $\hat{\sigma}_\rho^j$ is used from the regression on each joint j subsystem:

$$\boldsymbol{\sigma}_\rho^{j2} = \frac{\|\rho^j\|_{\min}^2}{r^j - p^j},$$

where $\|\rho^j\|_{\min}^2$, p^j, are the minimal norm error and the number of minimum parameters for each joint j subsystem, respectively. They are calculated with the Matlab 'economy size' QR decomposition of the joint j observation matrix, without calculating the LS solution.

The WLS solution $\hat{\boldsymbol{\chi}}_w$ minimizes the two norms of the vector of weighted errors $\boldsymbol{\rho}$:

$$\hat{\boldsymbol{\chi}}_w = \operatorname*{Arg\,min}_{\chi} [\boldsymbol{\rho}^T\mathbf{G}^T\mathbf{G}\boldsymbol{\rho}], \tag{7}$$

where $\hat{\chi}_{\mathrm{w}}$ and the corresponding standard deviations $\boldsymbol{\sigma}_{\hat{\chi}_{\mathrm{w}i}}$ are calculated as the LS solution of the system (4) weighted by $\mathbf{G}$:

$$\mathbf{y}_{\mathrm{w}} = \mathbf{W}_{\mathrm{w}}\boldsymbol{\chi} + \boldsymbol{\rho}_{\mathrm{w}}, \tag{8}$$

$$\mathbf{y}_{\mathrm{w}} = \mathbf{G}\mathbf{y}, \quad \mathbf{W}_{\mathrm{w}} = \mathbf{G}\mathbf{W}, \quad \boldsymbol{\rho}_{\mathrm{w}} = \mathbf{G}\boldsymbol{\rho}.$$

The unicity of $\hat{\chi}_{\mathrm{w}}$ depends on the observation matrix $\mathbf{W}_{\mathrm{w}}$ which can be numerically rank deficient depending on two origins:

- Structural rank deficiency which stands for any samples of $(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ in $\mathbf{W}_{\mathrm{w}}$. This is the structural parameter's identifiability problem which is solved using base parameters.
- Data rank deficiency due to a bad choice of noisy samples of $(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ in $\mathbf{W}_{\mathrm{w}}$. This is the problem of optimal measurement strategies which is solved using closed loop identification to track exciting trajectories.

In order to decrease the sensitivity of the LS solution of system (8) to errors in $\mathbf{y}_{\mathrm{w}}$ and $\mathbf{W}_{\mathrm{w}}$, the condition number of the observation matrix $\mathbf{W}_{\mathrm{w}}$ must be close to one with large singular values before computing $\hat{\chi}_{\mathrm{w}}$.

Cond($\mathbf{W}_{\mathrm{w}}$) = 1 means that all the standard deviation $\boldsymbol{\sigma}_{\hat{\chi}_{\mathrm{w}i}}$ are the same, that is to say, all the parameters are estimated with the same absolute accuracy. The drawback of this criterion for excitation is that the small parameters are poorly estimated because of a bad relative standard deviation (Eq. (10)).

If a priori knowledge $\bar{\chi}$ of χ is available, a better criterion is the condition number of the matrix $\mathbf{W}_{\mathrm{w}}$ weighted by $\bar{\chi}$ (Presse & Gautier, 1993):

$$\boldsymbol{\Phi} = \mathbf{W}_{\mathrm{w}}\,\mathbf{diag}(\bar{\chi}), \tag{9}$$

where $\mathbf{diag}(\bar{\chi})$ is a diagonal matrix with the coefficients of $\bar{\chi}$ on its diagonal.

Cond($\boldsymbol{\Phi}$) = 1 means that all the standard deviation divided by the a priori values, $\boldsymbol{\sigma}_{\hat{\chi}_{\mathrm{w}i}}/\bar{\chi}_i$, are the same. That is to say all the parameters are estimated with a good relative accuracy (Eq. (10)).

Exciting trajectories can be obtained by non-linear optimization of the coefficients of the trajectory generator as a polynomial interpolator (Gautier, 1992) or as a Fourier series (Swevers et al., 1997). In the following, it is supposed that this stage has been reached, that is to say $\mathbf{W}_{\mathrm{w}}$ is an $(r\times p)$ full rank and well conditioned matrix.

Standard deviations $\boldsymbol{\sigma}_{\hat{\chi}_{\mathrm{w}i}}$ are estimated considering the matrix $\mathbf{W}_{\mathrm{w}}$ to be a deterministic one. From Eq. (6), $\boldsymbol{\rho}_{\mathrm{w}}$ comes to be a zero mean additive independent noise such that

$$\mathbf{C}_{\boldsymbol{\rho}_{\mathrm{w}}\boldsymbol{\rho}_{\mathrm{w}}} = E(\boldsymbol{\rho}_{\mathrm{w}}^{\mathrm{T}}\boldsymbol{\rho}_{\mathrm{w}}) = \mathbf{I}_{\mathrm{r}}.$$

The covariance matrix of the estimation error and standard deviations can be calculated by

$$\mathbf{C}_{\hat{\chi}_{\mathrm{w}}\hat{\chi}_{\mathrm{w}}} = E\left[(\chi - \hat{\chi}_{\mathrm{w}})(\chi - \hat{\chi}_{\mathrm{w}})^{\mathrm{T}}\right] = (\mathbf{W}_{\mathrm{w}}^{\mathrm{T}}\mathbf{W}_{\mathrm{w}})^{-1},$$

where $\sigma^2_{\hat{\chi}_{\mathrm{w}i}} = \mathbf{C}_{\hat{\chi}_{\mathrm{w}}\hat{\chi}_{\mathrm{w}ii}}$ is the ith diagonal coefficient of $\mathbf{C}_{\hat{\chi}_{\mathrm{w}}\hat{\chi}_{\mathrm{w}}}$.

The relative standard deviation $\%\boldsymbol{\sigma}_{\hat{\chi}_{\mathrm{wr}i}}$ is given by

$$\%\boldsymbol{\sigma}_{\hat{\chi}_{\mathrm{wr}i}} = 100\,\frac{\boldsymbol{\sigma}_{\hat{\chi}_{\mathrm{wr}i}}}{\hat{\chi}_{\mathrm{w}i}}. \tag{10}$$

3. Extended Kalman filtering

3.1. Identification model

The state-space model is obtained from the inverse dynamic model Eq. (1) as follows:

$$\dot{\mathbf{x}} = \begin{bmatrix} \ddot{\mathbf{q}} \\ \dot{\mathbf{q}} \end{bmatrix} = \begin{bmatrix} \mathbf{M}^{-1}[\boldsymbol{\tau} - \mathbf{N}(\mathbf{q}, \dot{\mathbf{q}})] \\ \dot{\mathbf{q}} \end{bmatrix} = \mathbf{f}(\mathbf{x}, \boldsymbol{\tau}), \tag{11}$$

where $\mathbf{x} = [\,\dot{\mathbf{q}}^{\mathrm{T}} \quad \mathbf{q}^{\mathrm{T}}\,]^{\mathrm{T}}$ defines the state.

Computing the extended Kalman filter (Gautier et al., 1993; Guglielmi et al., 1987; Ljung, 1987) consists first in extending the state to include the model parameters to be identified and secondly in applying the Kalman filter equations (see Section 3.2).

Let us define the new extended state $\mathbf{z}$ including the parameters vector χ as

$$\mathbf{z} = \left[\,\dot{\mathbf{q}}^{\mathrm{T}} \quad \mathbf{q}^{\mathrm{T}} \quad \chi^{\mathrm{T}}\,\right]^{\mathrm{T}}.$$

Assuming that the parameters are stationary, the new augmented state space model is written as

$$\dot{\mathbf{z}} = \begin{bmatrix} \ddot{\mathbf{q}} \\ \dot{\mathbf{q}} \\ \dot{\chi} \end{bmatrix} = \begin{bmatrix} \mathbf{M}^{-1}[\boldsymbol{\tau} - \mathbf{N}(\mathbf{q}, \dot{\mathbf{q}})] \\ \dot{\mathbf{q}} \\ \mathbf{0}_{p\times 1} \end{bmatrix} \tag{12}$$

where $\mathbf{0}_{p\times 1}$ is a $(p\times 1)$ matrix of zeros.

3.2. Identification method

The first order discretization of Eq. (12) leads to

$$\mathbf{z}_{k+1} = \mathbf{f}(\mathbf{z}_k, \boldsymbol{\tau}_k) + \begin{bmatrix} \mathbf{v}_k \\ \mathbf{0}_{p\times 1} \end{bmatrix}, \tag{13}$$

$\mathbf{f}(\mathbf{z}_k, \boldsymbol{\tau}_k) = \mathbf{z}_k + \dot{\mathbf{z}}(\mathbf{z}_k, \boldsymbol{\tau}_k)\,\mathrm{d}t$, where $\mathrm{d}t$ is the sample rate, $\mathbf{v}_k$ is assumed to be a white noise sequence.

$$\mathbf{Q} = E\left(\begin{bmatrix} \mathbf{v}_k \\ \mathbf{0}_{p\times 1} \end{bmatrix}\begin{bmatrix} \mathbf{v}_k \\ \mathbf{0}_{p\times 1} \end{bmatrix}^{\mathrm{T}}\right)$$

is its covariance matrix.

The observation equations are given by

$$\mathbf{y}_k = \mathbf{q}_k = \mathbf{C}\mathbf{z}_k + \mathbf{w}_k, \quad \mathbf{C} = [\,\mathbf{0}_{n\times n} \quad \mathbf{I}_n \quad \mathbf{0}_{n\times p}\,],$$

where $\mathbf{w}_k$ is the measurement noise which is assumed to be a zero mean and independent sequence,

$$\mathbf{R} = E[\mathbf{w}_k\mathbf{w}_k^{\mathrm{T}}]$$

is its covariance matrix.

In order to achieve the estimation step and the calculation of the covariance matrices, the non-linear discrete equation (13) is linearized to the first order using a Taylor expansion around the estimate $\mathbf{z}_{k/k}$,

$$\mathbf{z}_{k+1} \cong \mathbf{f}(\mathbf{z}_{k/k}) + \mathbf{F}_z(\mathbf{z}_{k/k})(\mathbf{z}_k - \mathbf{z}_{k/k}) + \begin{bmatrix} \mathbf{v}_k \\ \mathbf{0}_{p\times 1} \end{bmatrix},$$

where $\mathbf{F}_z$ is the $(2n+p)\times(2n+p)$ Jacobian matrix of $\mathbf{f}$ with respect to $\mathbf{z}$.

The coefficients of $\mathbf{F}_z$ are defined as

$$F_{zij} = \frac{\partial f_i(\mathbf{z})}{\partial z_j}.$$

The optimal Kalman filtering is applied to the linearized system. It is composed of two steps

1. *Prediction step*:

$$\mathbf{z}_{k+1/k} = \mathbf{f}(\mathbf{z}_{k/k}),$$

$$\mathbf{P}_{k+1/k} = \mathbf{F}_z(\mathbf{z}_{k/k})\mathbf{P}_{k/k}\mathbf{F}_z(\mathbf{z}_{k/k})^{\mathrm{T}} + \mathbf{Q}.$$

2. *Estimation step*:

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1/k}\mathbf{C}^{\mathrm{T}}[\mathbf{C}\mathbf{P}_{k+1/k}\mathbf{C}^{\mathrm{T}} + \mathbf{R}]^{-1},$$

$$\mathbf{z}_{k+1/k+1} = \mathbf{z}_{k+1/k} + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \mathbf{C}\mathbf{z}_{k+1/k}),$$

$$\mathbf{P}_{k+1/k+1} = \mathbf{P}_{k+1/k} - \mathbf{K}_{k+1}\mathbf{C}\mathbf{P}_{k+1/k},$$

where $\mathbf{K}_{k+1}$ is the Kalman gain at time t_{k+1}, $\mathbf{z}_{k+1/k+1}$ the expected value of $\mathbf{z}_{k+1}$ given the $k+1$ measurements

$$\mathbf{z}_{k+1/k+1} = E(\mathbf{z}_{k+1}/y_i, \quad i = 1, \ldots, k+1),$$

$\mathbf{P}_{k+1/k}$ is the covariance matrix of the prediction error

$$\mathbf{P}_{k+1/k} = E[(\mathbf{z}_{k+1} - \mathbf{z}_{k+1/k})(\mathbf{z}_{k+1} - \mathbf{z}_{k+1/k})^{\mathrm{T}}/y_i, \quad i = 1, \ldots, k],$$

$\mathbf{P}_{k+1/k+1}$ is the covariance matrix of the estimation error

$$\mathbf{P}_{k+1/k+1} = E[(\mathbf{z}_{k+1} - \mathbf{z}_{k+1/k+1})(\mathbf{z}_{k+1} - \mathbf{z}_{k+1/k+1})^{\mathrm{T}}/y_i, \quad i = 1, \ldots, k+1].$$

In order to prevent some numerical problems, it is necessary to compute the algorithm by using a square root factorization as a $\mathbf{UDU}^{\mathrm{T}}$ decomposition which will ensure the covariance matrix remains a positive definite matrix (Bierman, 1977).

4. Description of the SCARA robot

The experimental comparison is carried out on a two-joints planar direct drive prototype robot manufactured in the laboratory (IRCCyN) (Figs. 1–3), without gravity effect. The description of the geometry of the robot uses

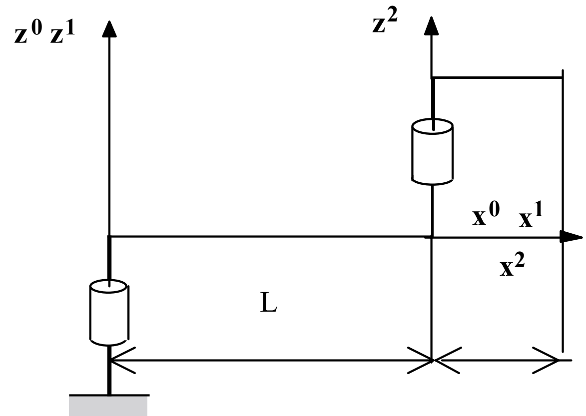

Fig. 1. SCARA robot.

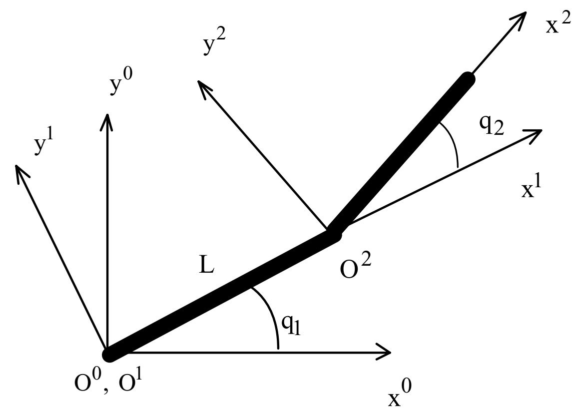

Fig. 2. Frames and joint variables.

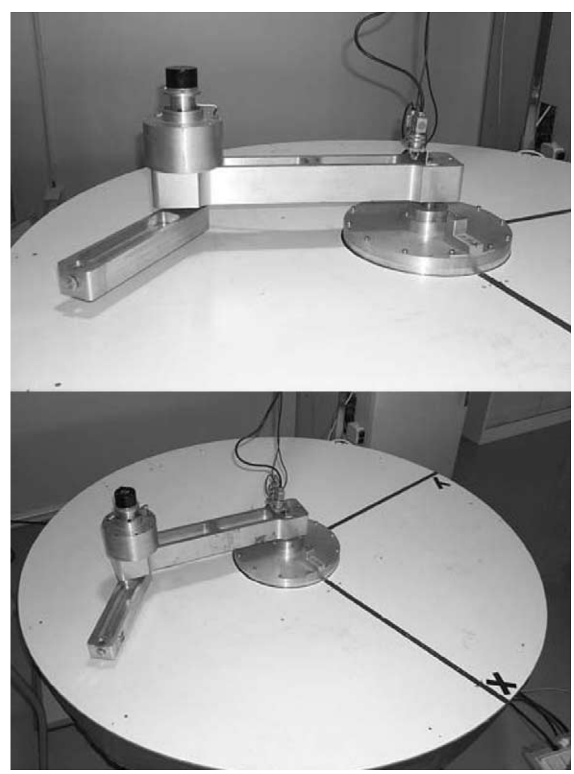
Fig. 3. Pictures of the SCARA.

the modified Denavit and Hartenberg notation (Khalil & Dombre, 1999).

The robot is directly driven by two DC permanent magnet motors supplied by PWM choppers.

The dynamic model depends on eight minimal dynamic parameters, including four friction parameters:

$$\chi = \begin{bmatrix} ZZR_1 & Fv_1 & Fs_1 & ZZ_2 & LMX_2 & LMY_2 & Fv_2 & Fs_2 \end{bmatrix}^{\mathrm{T}},$$

$$ZZR_1 = ZZ_1 + M_2L^2,$$

where L is the length of the first link, M_2 the mass of link 2, ZZ_1 and ZZ_2 the drive side moment of inertia of links 1 and 2, respectively, LMX_2 and LMY_2 the first moments of link 2 multiplied by the length L of link 1, Fv_1, Fs_1, Fv_2, Fs_2 are the viscous and coulomb friction parameters of links 1 and 2, respectively.

The inverse dynamic model used to compute WLS is written as Eq. (3):

$$\begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} \ddot{q}_1 & \dot{q}_1 & \mathrm{sign}(\dot{q}_1) & (\ddot{q}_1 + \ddot{q}_2) & \begin{pmatrix} (2\ddot{q}_1 + \ddot{q}_2)C_2 \\ -\dot{q}_2(2\dot{q}_1 + \dot{q}_2)S_2 \end{pmatrix} & \begin{pmatrix} -(2\ddot{q}_1 + \ddot{q}_2)S_2 \\ -\dot{q}_2(2\dot{q}_1 + \dot{q}_2)C_2 \end{pmatrix} & 0 & 0 \\ 0 & 0 & 0 & (\ddot{q}_1 + \ddot{q}_2) & (\ddot{q}_1C_2 + \dot{q}_1^2S_2) & (\dot{q}_1^2C_2 - \ddot{q}_1S_2) & \dot{q}_2 & \mathrm{sign}(\dot{q}_2) \end{bmatrix}\chi,$$

where $C_2 = \cos(q_2)$ and $S_2 = \sin(q_2)$.

The direct dynamic model necessary to compute the extended Kalman filtering algorithm is written as Eq. (11) with $\mathbf{q} = \begin{bmatrix} q_1 & q_2 \end{bmatrix}^{\mathrm{T}}$ and $\boldsymbol{\tau} = \begin{bmatrix} \tau_1 & \tau_2 \end{bmatrix}^{\mathrm{T}}$:

$$\begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} = \mathbf{M}^{-1} \begin{bmatrix} t_1 - F_{v1}\dot{q}_1 - F_{s1}\,\mathrm{sign}(\dot{q}_1) + LMX_2S_2\dot{q}_2(\dot{q}_2 + 2\dot{q}_1) + LMY_2C_2\dot{q}_2(\dot{q}_2 + 2\dot{q}_1) \\ t_2 - F_{v2}\dot{q}_2 - F_{s2}\,\mathrm{sign}(\dot{q}_2) - LMX_2S_2\dot{q}_1^2 - LMY_2C_2\dot{q}_1^2 \end{bmatrix},$$

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{12} & ZZ_2 \end{bmatrix},$$

$$M_{11} = ZZR_1 + ZZ_2 + 2LMX_2C_2 - 2LMY_2S_2,$$

$$M_{12} = ZZ_2 + LMX_2C_2 - LMY_2S_2.$$

The joint position $\mathbf{q}$ and the current reference $\mathbf{V}_T$ (the control input) are collected at a 100 Hz sample rate while the robot is tracking a fifth order polynomial trajectory. This trajectory has been calculated in order to obtain a good condition number $\mathrm{Cond}(\mathbf{W}_{\mathrm{w}}) = 290$ and $\mathrm{Cond}(\boldsymbol{\Phi}) = 100$. This means that it is an exciting trajectory taking the whole trajectory all over at the time of the test. Both methods are performed in a closed loop identification scheme (simply joint PD control), using the same data $\mathbf{q}$ and $\boldsymbol{\tau}$, where each torque $\boldsymbol{\tau}_j$ is calculated as

$$\boldsymbol{\tau}_j = G_{Tj}V_{Tj},$$

where G_{Tj} is the drive chain gain which is considered as a constant in the frequency range of the robot dynamics. Fig. 4 presents the torque of motors 1 and 2.

5. Experimental results

5.1. WLS estimation

Calculation of the WLS solution of Eq. (4) (or the LS solution of Eq. (8)) from noisy discrete measurements or estimations of $(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}, \boldsymbol{\tau})$ may lead to bias because $\mathbf{W}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ and $\mathbf{y}(\boldsymbol{\tau})$ may be non-independent random matrices. Therefore, it is very important to decrease any perturbation in $\mathbf{W}$ and $\mathbf{y}$ before computing the WLS solution. The joint velocities and accelerations are estimated with bandpass filtering of $\mathbf{q}$ using a non-causal zero-phase digital filter with flat amplitude characteristic (lowpass butterworth filter in both the forward and reverse direction as given by the filtfilt procedure from Matlab, and central difference algorithm for derivatives) (Gautier, 1997). It is very important to avoid distortion in $\mathbf{W}$ due to the filtering of $\mathbf{q}$ and its derivatives. As a second step, the vector $\mathbf{y}$ and each column of $\mathbf{W}$ are lowpass filtered and decimated. Due to the linearity of Eq. (4), the WLS is not sensitive to the distortion introduced by this filtering.

The elapsed time for the computation is about 3 s on a Pentium II 233 MHz PC. Table 1 compares WLS estimation obtained:

1. *Without any lowpass filtering*: see the biais on ZZR_1, Fs_1, LMX_2, Fv_2, Fs_2.
2. *With only forward lowpass filtering*: see the biais on ZZR_1 and Fv_2.
3. *With full filtering*: All the parameters have significant values close to a priori values $\bar{\boldsymbol{\chi}}$, with small relative standard deviations, except for LMY_2 and Fv_1. They are too small to have a significant contribution in $\mathbf{y}_{\mathrm{w}}$.

They can be cancelled from the dynamic model to get only six essential parameters.

Remarks.

- A priori values $\bar{\chi}$ have been calculated from measurements on the disassembled links or special tests moving one link at a time. They are given in Tables 1 and 2, to be compared with the identified parameters. They have been used to optimize the trajectory, Eq. (9).
- The LS solution of Eq. (8) is calculated with a global QR factorization of $\mathbf{W}_w$ (left matrix divide in Matlab). This is a known robust and fast method provided there is a vector form calculation and no disk swap. The recursive least squares (RLS) is not needed for batch processing and it takes much more time because of the loop form calculation in Matlab. However, the RLS results are given in Figs. 13–28 for comparison with EKF convergence. In order to avoid large and non-significant transients during the first samples $j<p$, RLS is initialized with the square solution using the first p samples in Eq. (8).

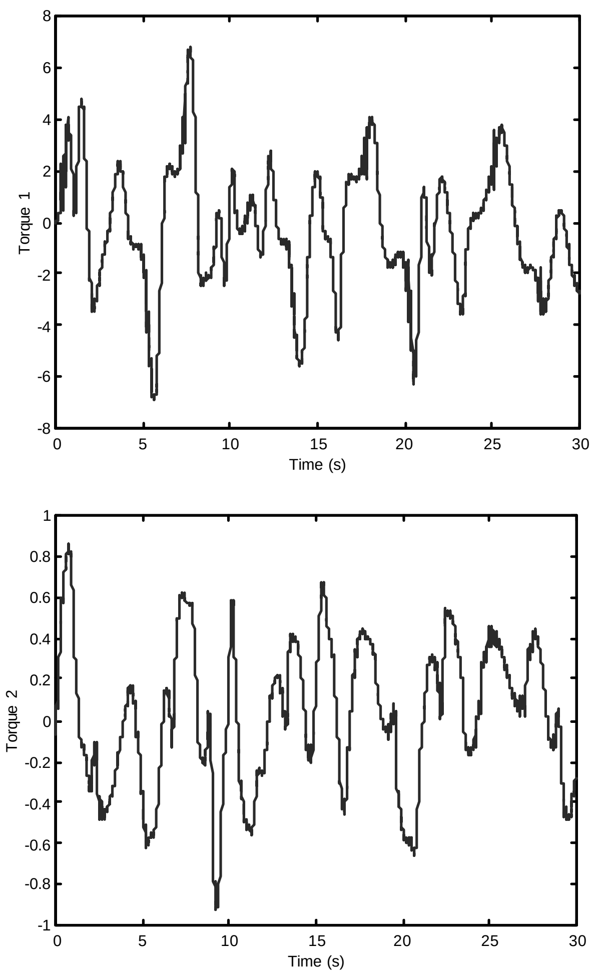

Fig. 4. Motor torque.

5.2. EKF estimation

Three experimental cases are introduced depending on the accuracy of the a priori knowledge on parameter values. The identification is performed using only the torque and the joint position measurement. Joint velocities are estimated through the Kalman filter.

1. *Without knowledge*: $\hat{\boldsymbol{\chi}}_0 = \mathbf{0}_{8\times 1}$ and $\mathbf{P}_{0/0} = 10^5\mathbf{I}_{12}$, does not work because of numerical problems.
2. *Rough a priori knowledge*: The sign and the size of each parameter are assumed to be known from manufacturer's data

$$\hat{\chi}_0 = \begin{bmatrix} 1 & 10^{-3} & 10^{-1} & 10^{-1} & 10^{-1} & 10^{-3} & 10^{-3} & 10^{-1} \end{bmatrix}^{\mathrm{T}},$$

$$\mathbf{P}_{0/0} = \mathbf{diag}([\,10 \quad 10 \quad 10 \quad 10 \quad 10 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1\,]).$$

The elapsed time for the computation is about 170 s on a Pentium II 233 MHz PC. The variance matrix for the state (position and velocity) noise and the measurement noise are given by

$$\mathbf{R} = \mathbf{diag}([\,10^{-3} \quad 10^{-2}\,]), \quad \mathbf{Q} = 10^{-2}\mathbf{I}_4.$$

Table 1
WLS estimation

		WLS (1)		WLS (2)		WLS (3)	
	$\bar{\chi}$	$\hat{\chi}$	$\%\sigma_{\hat{\chi}}$	$\hat{\chi}$	$\%\sigma_{\hat{\chi}}$	$\hat{\chi}$	$\%\sigma_{\hat{\chi}}$
ZZR_1	3.42	0.51	4.8	3.26	1.2	3.47	0.8
Fv_1	0.07	−0.66	24	−0.2	60	0.3	25
Fs_1	0.58	0.81	9.4	0.55	9.6	0.4	9
ZZ_2	0.064	0.056	0.73	0.064	0.8	0.063	0.45
LMX_2	0.131	0.042	4	0.13	1.7	0.125	1.2
LMY_2	0	0.002	65	0.004	60	0.003	49
Fv_2	0.015	0.021	6	−0.006	22	0.014	6
Fs_2	0.156	0.13	2.5	0.17	2.3	0.13	1.7

Table 2
EKF estimation

	$\bar{\chi}$	EKF (2) $\hat{\chi}$	EKF (2) $\%\sigma_{\hat{\chi}}$	EKF (3) $\hat{\chi}$	EKF (3) $\%\sigma_{\hat{\chi}}$	WLS $\hat{\chi}$	WLS $\%\sigma_{\hat{\chi}}$
ZZR_1	3.42	3.21	0.17	3.31	0.2	3.47	0.8
Fv_1	0.07	−0.95	−1.67	0.6	3.6	0.3	25
Fs_1	0.58	1.28	0.53	0.58	1.4	0.4	9
ZZ_2	0.064	0.06	0.14	0.061	0.11	0.063	0.45
LMX_2	0.131	0.123	0.49	0.14	0.39	0.125	1.2
LMY_2	0	0.064	0.69	0.03	1.05	0.003	49
Fv_2	0.015	0.0159	1.57	0.012	1.12	0.014	6
Fs_2	0.156	0.127	0.64	0.13	0.34	0.13	1.7

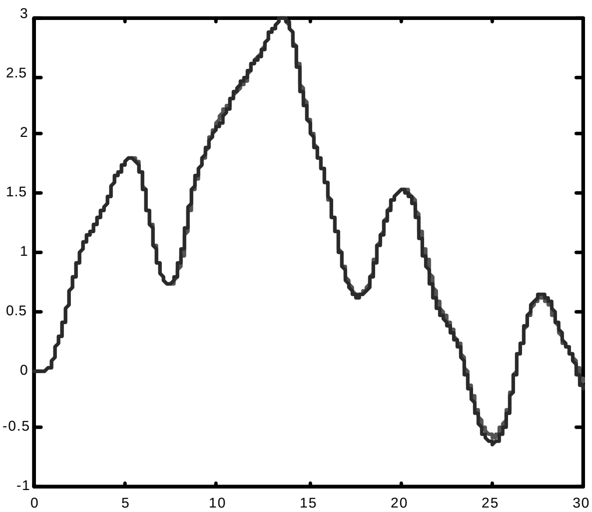
Fig. 5. Joint 1 position.

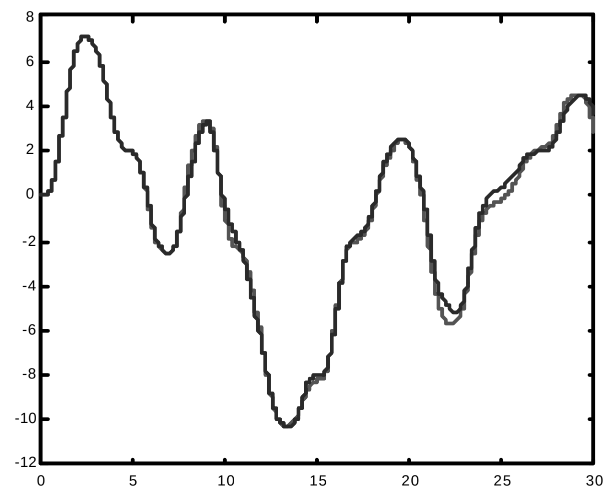
Fig. 7. Joint 2 position.

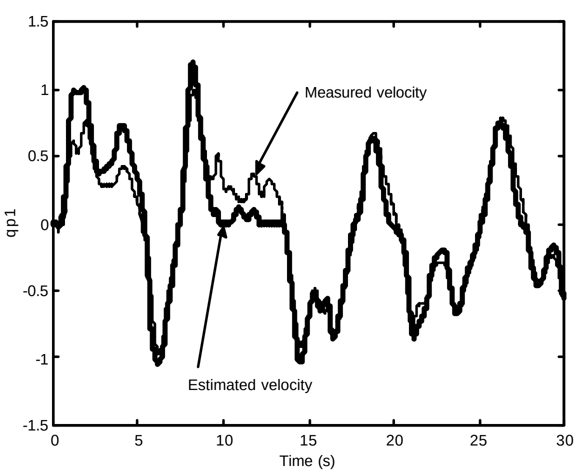

Fig. 6. Joint 1 velocity.

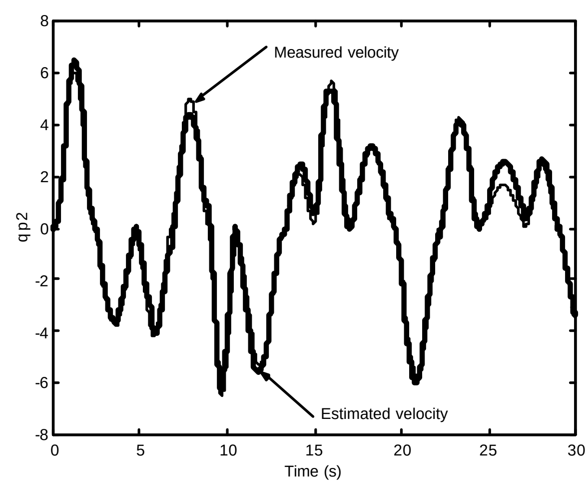

Fig. 8. Joint 2 velocity.

Figs. 5–8 give the positions and velocities estimated with EKF and calculated by band pass filtering (named measured). See Section 6 to explain the differences.

3. *Good a priori knowledge*: The EKF algorithm is initialized with WLS results. The parameters (ZZR_1, Fs_1, ZZ_2, LMX_2, Fv_2, Fs_2) which are well identified, are locked with a small initial variance (10^{-3}) while

Fv_1 and LMY_2 are free (initial variance 1)

$$\hat{\chi}_0 = [3.47 \quad 10^{-3} \quad 0.4 \quad 0.063 \quad 0.125 \quad 10^{-3} \quad 0.014 \quad 0.13]^{\mathrm{T}},$$

$$\mathbf{P}_{0/0} = \mathbf{diag}([10 \quad 10 \quad 10 \quad 10 \quad 10^{-3} \quad 1 \quad 10^{-3} \quad 10^{-3} \quad 10^{-3} \quad 1 \quad 10^{-3} \quad 10^{-3}]).$$

The elapsed time for the computation is about 170 s.

The variance matrix for the state (position and velocity) noise and the measurement noise are given as

$$\mathbf{R} = \mathbf{diag}([10^{-4} \quad 10^{-2}]), \quad \mathbf{Q} = 10^{-4}\mathbf{I}_4.$$

Figs. 9–12 represent the measured and estimated signals for joint position and velocity.

The reader will find in the Appendix the estimated parameters with respect to the time in case of EKF(2) (Figs. 13–20) with rough a priori knowledge and EKF(3) (Figs. 21–28) with good a priori knowledge. It can be seen that the recursive WLS estimation convergence is faster with a better steady state.

6. Discussion

Some major points have to be highlighted

6.1. Concerning EKF algorithm

1. The measured position is used roughly without particular treatment as for the WLS case.
2. The results are very sensitive with respect to initial values, and good a priori knowledge is highly recommended. It can be obtained as the WLS estimation.

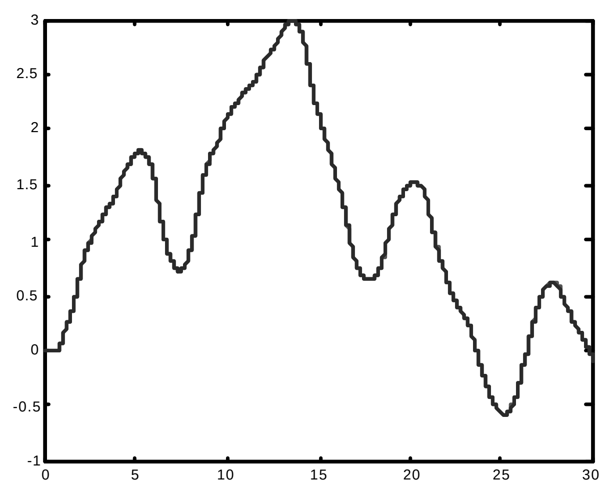

Fig. 9. Joint 1 position.

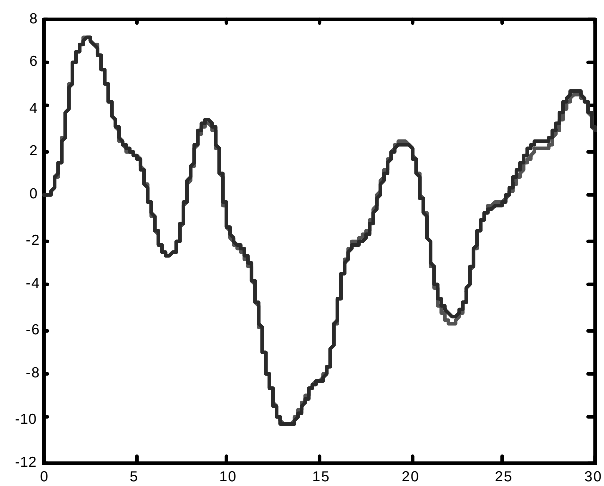

Fig. 11. Joint 2 position.

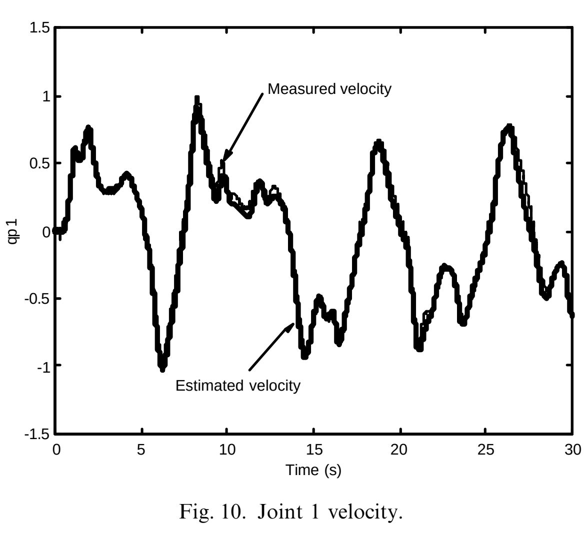

Fig. 10. Joint 1 velocity.

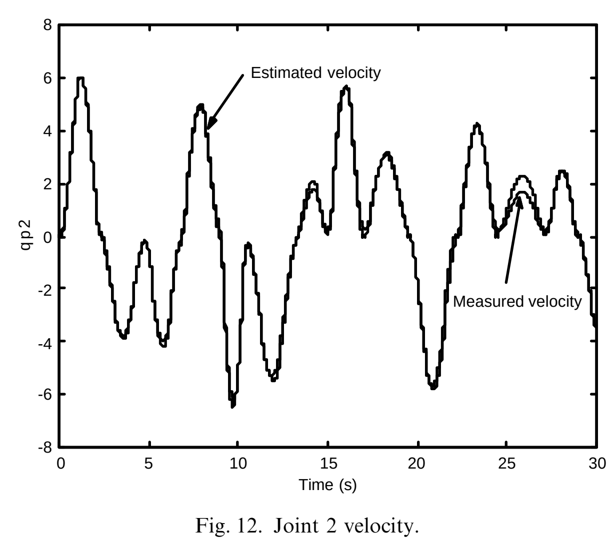

Fig. 12. Joint 2 velocity.

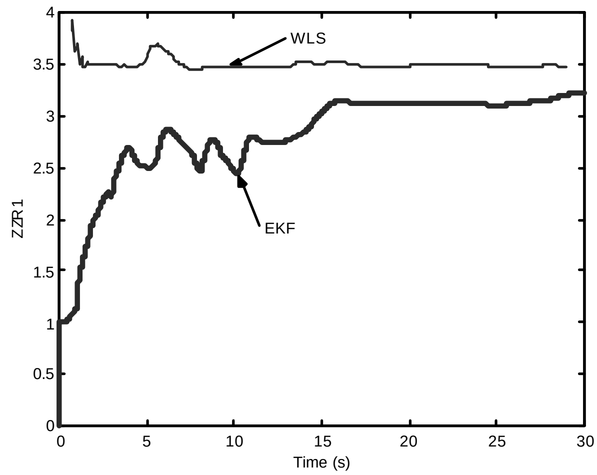

Fig. 13. ZZR_1.

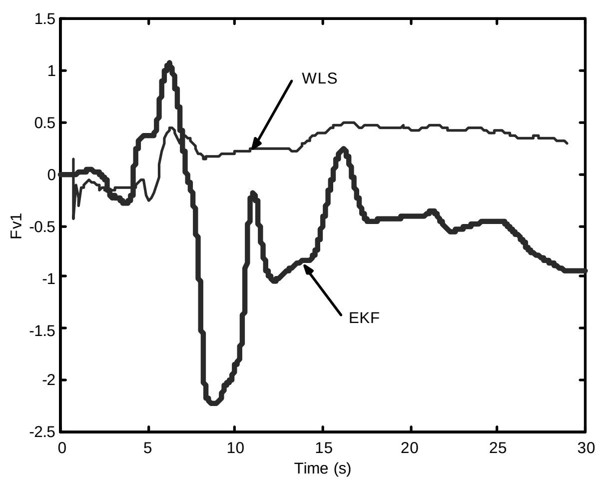

Fig. 14. Fv_1.

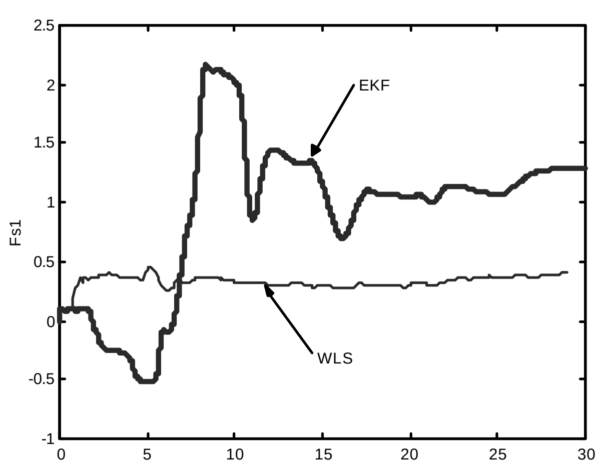

Fig. 15. Fs_1.

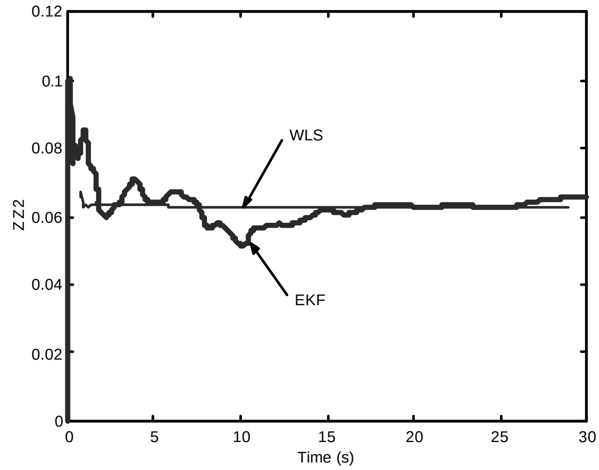

Fig. 16. ZZ_2.

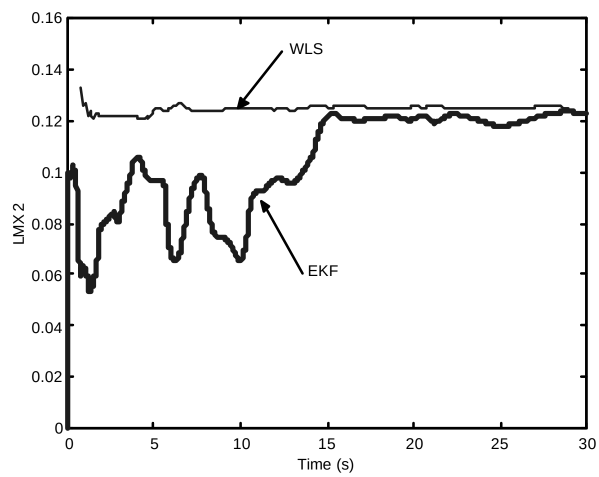

Fig. 17. LMX_2.

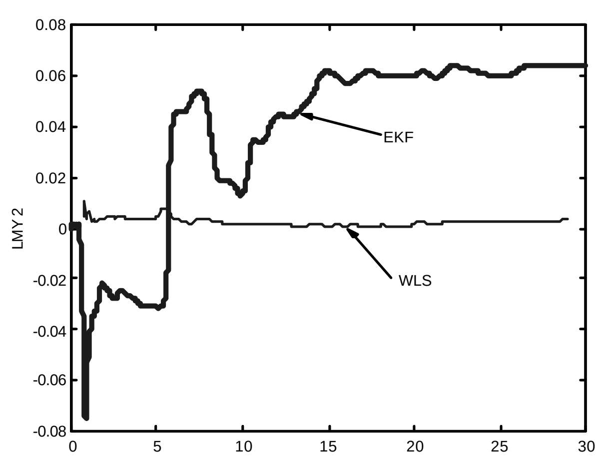

Fig. 18. LMY_2.

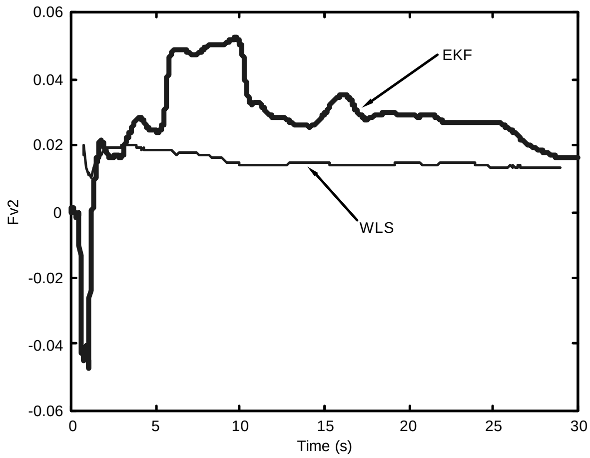

Fig. 19. Fv_2.

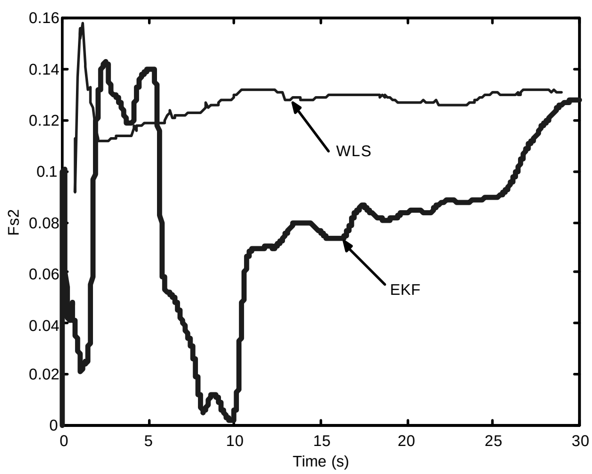

Fig. 20. Fs_2.

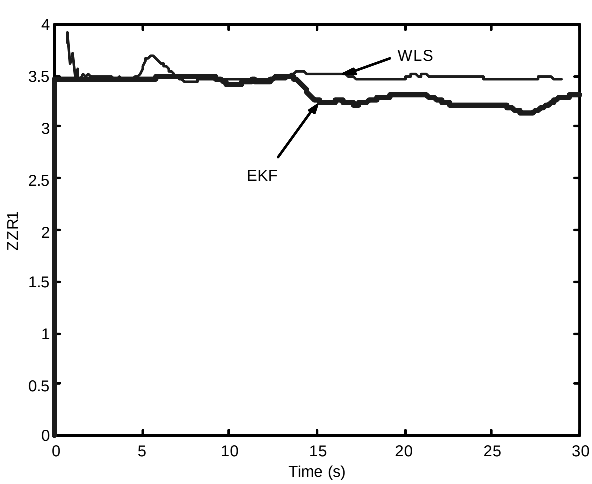

Fig. 21. ZZR_1.

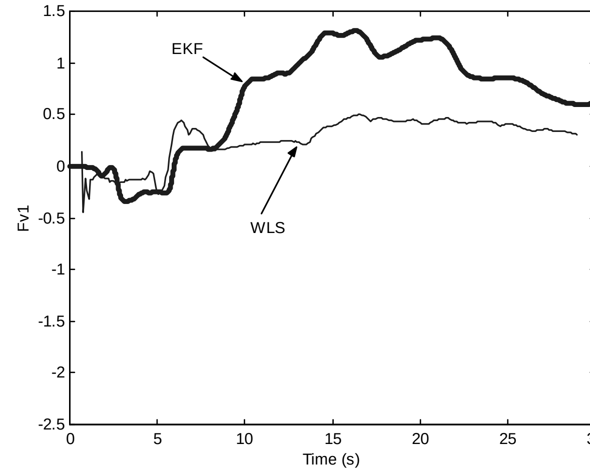

Fig. 22. Fv_1.

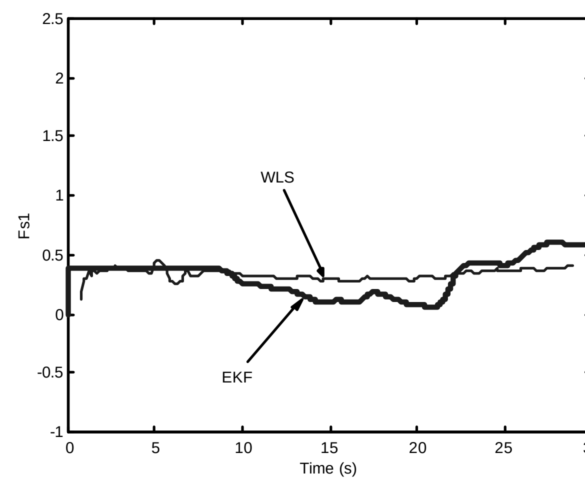

Fig. 23. Fs_1.

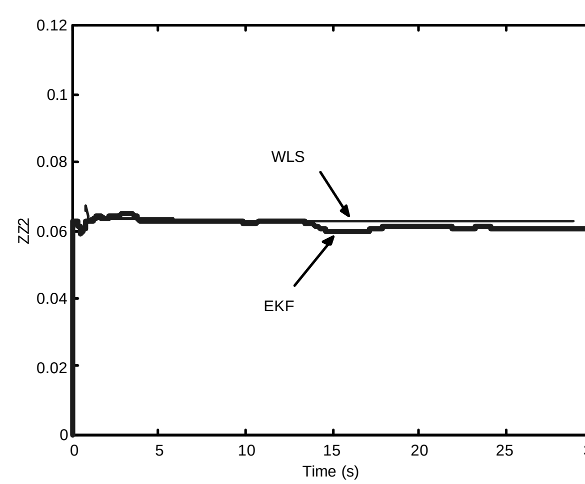

Fig. 24. ZZ_2.

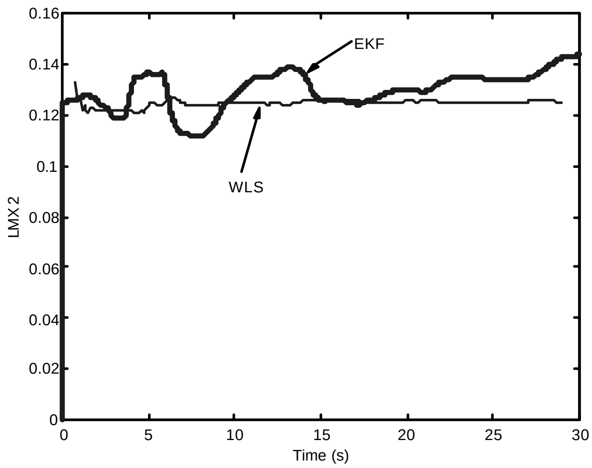

Fig. 25. LMX_2.

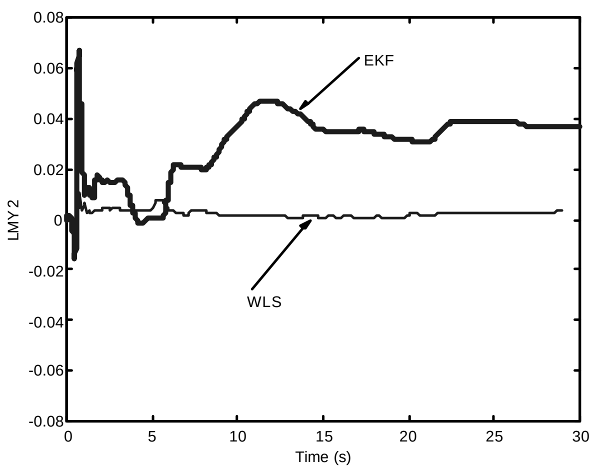

Fig. 26. LMY_2.

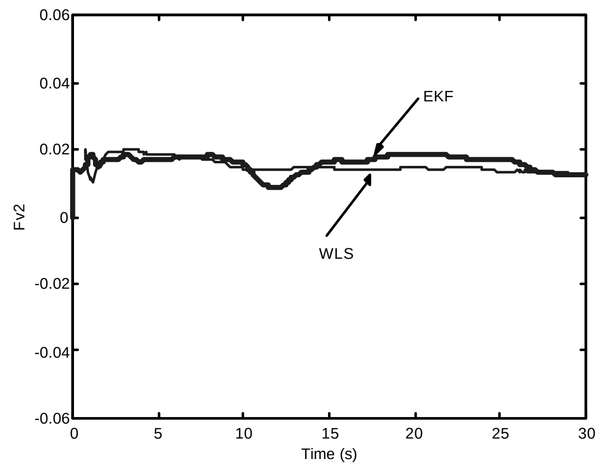

Fig. 27. Fv_2.

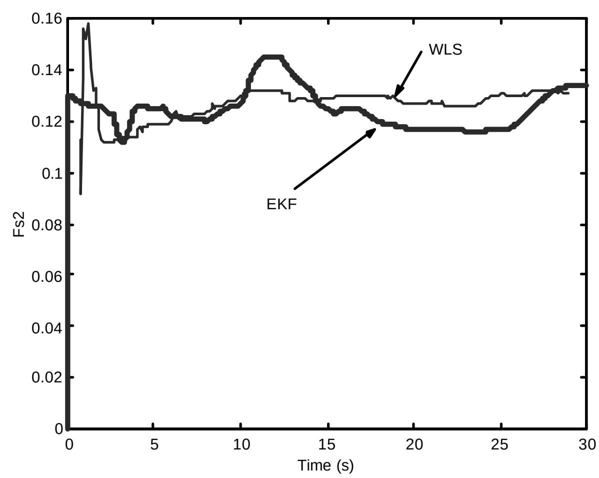

Fig. 28. Fs_2.

3. Velocities and parameters are estimated through the same model which is a very significant result of the comparative study, but not an advantage. The EKF is an on line state observer which is useful for control, but it does not take advantage of the off line identification such as the cancellation of the phase distortion.
4. The calculations of the direct dynamic model and its Jacobian matrix w.r.t. the extended state is tedious and lead to a time consuming algorithm using symbolic calculation software.
5. The recursive algorithm does not suit the vector form calculation. Then it takes a very long computing time using software such as Matlab for large amount of data, even when using the compiled version of the Matlab script.

6.2. Concerning WLS algorithm

1. The deterministic forward and reverse pass band filtering does not introduce neither phase distortion nor errors on the estimated derivatives due to the dynamic modeling errors as for EKF (see Figs. 6 and 10). However, the user has to take care to use filters without any distortion in the frequency range of the closed loop dynamics of the robot (see Table 1).
2. With the WLS method, the uncertainties on the measurements and the model are taken into account in only one global residual. The estimation of standard deviation is based on simple statistical hypotheses considering additive noise on the measurements (Gautier, 1997). If the standard deviation is too large, it may come from the modeling errors or not sufficient exciting trajectories compared with the measurements perturbations (e.g. Fv_1 and LMY_2 have very small contribution in $\mathbf{y}_w$).
3. Due to the linear model expression, the automatic calculation of the identification model is easier and faster.
4. A priori knowledge is not needed.

7. Conclusion

This paper investigates theoretical and experimental comparison of EKF and WLS estimation applied on a 2 d.o.f. SCARA robot. The major results are: EKF algorithm estimates both the velocities and the parameters while WLS estimation needs the joint velocity and acceleration to be calculated separately by pass band filtering. However, it does not appear to be an advantage for EKF. Estimations of the parameters are very close for both methods, but EKF algorithm is very sensitive to the initial conditions and the convergence speed is slower. Moreover, recursive calculations are time consuming, and symbolic calculation of the Jacobian matrix is very tedious for the EKF method.

The conclusion is that the WLS method with the inverse dynamic model appears to be better than EKF for off line identification.

Future work will concern the analysis of the on line behavior with a priori knowledge given by WLS and parameter tracking with EKF algorithm.

Appendix A

Figs. 13–20 present the estimated parameters with rough a priori knowledge of EKF(2) and Figs. 21–28 present the estimated parameters with good a priori knowledge of EKF(3).

References

Bierman, G. J. (1977). *Factorization methods for discrete sequential estimation*. London: Academic Press.

Califiore, G., & Indri, M. (1999). *Identification of a robot dynamic model: Experiment design and parameter estimation*. International Summer School on Modern Control, Rome.

Canudas de Wit, C., Siciliano, B., & Bastin, G. (Eds), (1996). *Theory of robot control*. London: Springer.

Gautier, M. (1986). Identification of robot dynamics. *Proceedings of the IFAC symposium: Theory of robots*, Vienna, Autriche (pp. 351–356).

Gautier, M. (1990). *Contribution à la modélisation et à l'identification des robots*. Thèse d'état, Université de Nantes, ENSM Nantes.

Gautier, M. (1991). Numerical calculation of the base inertial parameters. *Journal of Robotics Systems*, *8*(4), 485–506.

Gautier, M. (1992). Optimal motion planning for robot's inertial parameters identification. *Proceedings of the 31st CDC*, Tucson (pp. 70–73).

Gautier, M. (1997). Dynamic identification of robots with power model. *International conference on robotics and automation*, Albuquerque, USA (pp. 1922–1927).

Gautier, M., & Khalil, W. (1988). On the identification of the inertial parameters of robots. *Proceedings of the 27th CDC* (pp. 2264–2269).

Gautier, M., & Khalil, W. (1990). Direct calculation of minimum inertial parameters of serial robots. *IEEE Transactions on Robotics and Automation*, *6*(3), 368–373.

Gautier, M., Janin, C., & Presse, C. (1993). Dynamic identification of robots using least squares and extended Kalman filtering methods. *Proceedings of ECC'93* (pp. 2291–2297).

Guglielmi, M., Jonker, E., & Piasco, J. M. (1987) Modelling and identification of a two degrees of freedom SCARA robot using extended Kalman filtering. *International conference on advanced robotics*, Paris (pp. 137–148).

Khalil, W., & Creusot, D. (1997). SYMORO+: A system for the symbolic modelling of robots. *Robotica*, *15*, 153–161.

Khalil, W., & Dombre, E. (1999). *Modélisation, identification et commande des robots*. Hermès, Paris.

Kozlowski, K. (1998). *Modelling and identification in robotics*. London: Springer.

Ljung, L. (1987). *System identification, theory for the user*. Englewood Cliffs, NJ: Prentice-Hall.

Mayeda, H., Yoshida, K., & Osuka, K. (1990). Base parameters of manipulator dynamic models. *IEEE Transactions On Robotics and Automation*, *RA-6*(3), 312–321.

Presse, C., & Gautier M., (1993). New criteria of exciting trajectories for robot identification. *IEEE international conference on robotics and automation*, Atlanta, Georgia, USA (pp. 907–912).

Swevers, J., Ganseman, C., Tükel, D. B., Schutter, J. D., & Van Brussel, H. (1997). Optimal robot excitation and Identification. *IEEE Transactions on Robotics and Automation*, *13*(5), 730–740.