2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Wavelet Networks for Estimation of Coupled Friction in Robotic Manipulators

Edvard Naerum*†, Jordi Cornellà*†and Ole Jakob Elle*
*The Interventional Centre, Rikshospitalet University Hospital, Oslo, Norway
†Faculty of Medicine, University of Oslo, Norway
E-mail: edvard.narum@medisin.uio.no

***Abstract*— A wavelet network (WN) friction model has been developed for robots where the friction is coupled, such that it is a function of the velocity of multiple joints. Wavelets have the ability to estimate random friction maps without any prior modeling while preserving linearity in the model parameters. The WN friction model was compared against the Coulomb+viscous (CV) model through experiments with the PHANTOM Omni haptic device (SensAble Technologies, MA, USA); however, the theory is valid for any serial-chain robotic manipulator. Ability to estimate applied motor torques was used as the performance metric, quantified using relative RMS values. During training of the WN model it outperformed the CV model in all cases, with an improvement in relative RMS ranging from 0.4 to 7.5 percentage points, illustrating the potential of the WN friction model. However, during testing of the WN model on an independent data set results were mixed, highlighting the challenge of achieving sufficient training.**

I. Introduction

When building the dynamic model of a robot the joint friction presents a difficult challenge. Our goal is to utilize the knowledge of the dynamic model to estimate the contact forces when a robot is interacting with soft tissue during teleoperated surgery. In this case the robot will operate at low velocities, where friction effects are dominant. The most common friction model used in engineering is the Coulomb+viscous (CV) model [1], [2]. Several useful friction models also exist that can capture more complex phenomena such as hysteresis, pre-sliding displacement, varying break-away force and the Stribeck effect [3], [4], [5], [6]. When the friction in a joint of the robot is coupled, in the sense that it depends on the velocity in multiple joints and/or actuators, friction modeling is complicated further. Prisco and Bergamasco [7] modeled viscous friction in a class of tendon-driven manipulators by employing a dissipation function. Kobayashi and Ozawa [8] proposed to use neural networks to model tendon-driven mechanisms, thus also attempting to capture the inherent joint friction.

The ideal dynamic equation of a robot—computed using for instance the Lagrangian—is linear in the model parameters. For that reason it is desirable to use a friction model that preserves this property, like the CV model, to simplify analysis and implementation. De Wit et al. [9] proposed a linear-in-parameter extension to the CV model to approximate the Stribeck effect. Jatta et al. [10] performed contour tracking with a robot whose friction was modeled with a polynomial function. Non-classical techniques like wavelets, neural networks and fuzzy systems have also been proposed for friction modeling. Huang et al. [11] performed simulations with a single mass system where the Stribeck effect was compensated for using a radial basis function (RBF) network; Du and Nair [12] presented a method to compensate for friction in a DC motor at low velocities using wavelets; Santibañez et al. [13] proposed a fuzzy controller where the steady-state position error could be arbitrarily reduced even in the presence of static friction; and Gomes et al. [14] applied a neuro-fuzzy system to compensate for friction where the friction was allowed to drift over time.

A finite-size expansion (network) of wavelets can approximate random non-linear functions such as a friction map to a desired accuracy, as long as one can assume that the friction is sufficiently described as a static function of one or more variables, such as velocity. In addition to being linear in the parameters a wavelet network (WN) friction model does not make a priori assumptions about the shape of the friction map. In this paper a general WN friction model is developed for a robot with d degrees of freedom (DOF), where the friction in each joint can be a function of the velocity of one or more joints. The best fit wavelets are automatically selected with a thresholding procedure, and a Lyapunov approach is used to prove the stability and convergence properties of the parameter adaptation law. The model is compared against the CV model, and it is hypothesized that because of its function approximation ability the WN model will perform better. It is not attempted to compete with the more advanced dynamic friction models, the intention is rather to build a model which is conceptually as simple as the CV model, but with improved performance. The experiments are performed using the PHANTOM Omni haptic device from SensAble Technologies, Massachusetts, USA.

II. Problem Statement

The dynamics of a d-DOF robot can be described as

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{N}(\boldsymbol{q}) + \boldsymbol{\tau}_f(\dot{\boldsymbol{q}}) = \boldsymbol{\tau} \qquad (1)$$

where $\boldsymbol{q} \in \mathbb{R}^d$ is the angle vector, $\boldsymbol{M} \in \mathbb{R}^{d\times d}$ is the manipulator inertia matrix, $\boldsymbol{C} \in \mathbb{R}^{d\times d}$ is the Coriolis matrix, $\boldsymbol{N} \in \mathbb{R}^d$ is the gravity vector, $\boldsymbol{\tau}_f \in \mathbb{R}^d$ is the joint friction vector and $\boldsymbol{\tau} \in \mathbb{R}^d$ is the motor torque vector. The problem to be solved can be stated as follows.

Friction estimation problem: Given $\boldsymbol{M}$, $\boldsymbol{C}$, $\boldsymbol{N}$, $\boldsymbol{q}$ and $\boldsymbol{\tau}$, compute an estimate of the robot's joint friction $\boldsymbol{\tau}_f$. The

friction model should be linear in the parameters, and it should be able to take coupled friction into account.

Two main (and independent) assumptions are made.

Assumption 1: The friction in the joints of the robot is sufficiently described with a static map from velocity to force. Dynamic friction phenomena are neglected, in order to have a friction model which is linear in its parameters.

Assumption 2: The angular velocity $\dot{\boldsymbol{q}}$ remains small. In particular, it is confined to a bounded subset $\mathcal{X}_r \subset \mathbb{R}^d$, and the approximation $\dot{\boldsymbol{M}} \approx \boldsymbol{0}$ can be justified.

Initially the matrices $\boldsymbol{M}$, $\boldsymbol{C}$ and $\boldsymbol{N}$ are not given either. Hence, for the purpose of parameter identification (1) is rewritten (without the friction term $\boldsymbol{\tau}_f$) so that the unknown parameters occur linearly as:

$$\boldsymbol{Y}(\ddot{\boldsymbol{q}}, \dot{\boldsymbol{q}}, \boldsymbol{q})\boldsymbol{\alpha} = \boldsymbol{\tau} \tag{2}$$

where $\boldsymbol{Y} \in \mathbb{R}^{d \times p}$ is the regressor matrix, $\boldsymbol{\alpha} \in \mathbb{R}^p$ is the parameter vector, and p is the number of unknown parameters. Acceleration measurements are often not available and must be obtained via numerical differentiation of the velocity, thereby producing noisy signals. In order to eliminate the dependency of acceleration from the regressor matrix, both sides of (2) are run through a first order low-pass filter $\omega/(s+\omega)$ [15]. The dynamic equation becomes

$$\boldsymbol{Y}_L(\dot{\boldsymbol{q}}, \boldsymbol{q})\boldsymbol{\alpha} = \boldsymbol{\tau}_L \tag{3}$$

where $\boldsymbol{Y}_L$ and $\boldsymbol{\tau}_L$ are the filtered versions of $\boldsymbol{Y}$ and $\boldsymbol{\tau}$, respectively. A linear optimization method can be used with (3) to identify the values of the elements in $\boldsymbol{\alpha}$.

III. Friction Estimation with Wavelet Networks

Below, a friction model based on wavelet theory is presented as a solution to the friction estimation problem stated in Section II. First, wavelets are presented as a general tool for function reconstruction, before adapting the general theory to the case of friction estimation.

A. Wavelets for Function Reconstruction

The discrete wavelet transform of a function $f(x) \in L^2(\mathbb{R})$ (f is square-integrable) is defined by

$$W_{m,n} = \langle f, \psi_{m,n} \rangle = a_0^{-m/2} \int_{-\infty}^{\infty} f(x)\psi(a_0^{-m}x - nb_0)dx$$

where $m, n \in \mathbb{Z}$, $a_0 > 1$ and $b_0 > 0$. The functions

$$\psi_{m,n}(x) = a_0^{-m/2}\psi(a_0^{-m}x - nb_0) \tag{4}$$

are called wavelets, and they are all generated by dilating (changing m) and translating (changing n) a mother wavelet $\psi(x)$. Wavelet theory states that f can be exactly reconstructed by the infinite expansion

$$f(x) = \sum_{m,n} c_{m,n}\psi_{m,n}(x) \tag{5}$$

under the condition that the $\psi_{m,n}$ constitute a frame, and where the parameters $c_{m,n}$ represent the wavelet transforms of the dual frame. The notion of wavelet frames is not explained in this paper, but for a comprehensive discussion of wavelet theory see for example Daubechies [16].

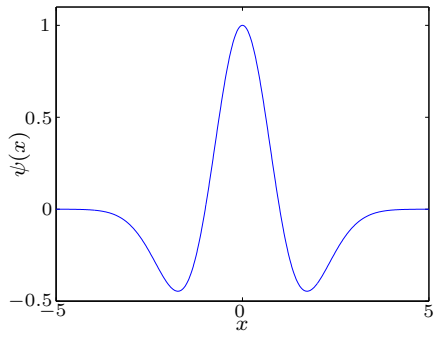

Fig. 1. The Mexican hat wavelet.

The Mexican hat function, shown in Fig. 1, is used as the mother wavelet in the proposed friction estimation algorithm. Mathematically it is given by

$$\psi(x) = (1 - x^2)e^{-x^2/2},$$

and with $a_0 = 2$ and $b_0 = 1$ the corresponding $\psi_{m,n}$ satisfy the requirement for a frame. For higher dimensions where $\boldsymbol{x} \in \mathbb{R}^d$ the wavelets (4) become

$$\psi_{m,\boldsymbol{n}}(\boldsymbol{x}) = a_0^{-dm/2}\psi(a_0^{-m}\boldsymbol{x} - \boldsymbol{n}b_0);\ m \in \mathbb{Z}, \boldsymbol{n} \in \mathbb{Z}^d.$$

The Mexican hat is extended by inserting the norm of $\boldsymbol{x}$, like

$$\psi(\boldsymbol{x}) = (1 - \|\boldsymbol{x}\|^2)e^{-\|\boldsymbol{x}\|^2/2}.$$

Notice that the dilated wavelet becomes narrower or more 'high frequent' as m decreases. Intuitively high frequent wavelets can be used where f is changing rapidly, while low frequent wavelets can be used where f is changing more slowly. For friction estimation it is possible to think that it is best to use an odd mother wavelet since friction is discontinuous around zero. However, high frequency Mexican hats can approximate this discontinuity while low frequency ones can approximate smoother regions of the friction map, where it is logical to use an even mother wavelet.

Wavelet theory requires that the function to be reconstructed is in L^2. While in general this is not the case with the friction $\boldsymbol{\tau}_f$, it can be said to be in L^2 by using Assumption 2. Friction is only approximated within $\mathcal{X}_r$. Outside $\mathcal{X}_r$ friction is regarded to be zero, even though that is not the case. It does not matter as long as the velocity of the robot never leaves $\mathcal{X}_r$.

B. Wavelet Network Friction Model

With wavelets no modeling is required. One only needs to identify, for each joint, which velocities affect the friction. The proposed friction model uses the multi-dimensional version of (5), but with a *finite* number of elements. One expansion is used for each joint, for a total of d expansions.

Let the friction in joint i be a function of d_i velocities, where $d_i \in \mathbb{Z}$ and $1 \leq d_i \leq d$. Furthermore let $\mathcal{K}_i \subset \mathbb{Z}^{1+d_i}$

be the index set containing the K_i indices $k_i = [m_i, \boldsymbol{n}_i]^T$ of the wavelets that may *potentially* be included in the expansion for joint i (the selection of the appropriate wavelets to be included in the expansions is explained in Section III-D). For the moment assume that all wavelets in $\mathcal{K}_i$ are included. The friction estimate $\hat{\tau}_{f,i}$ can then be written as

$$\hat{\tau}_{f,i}(\dot{\boldsymbol{q}}_{aff,i}) = \sum_{k_i \in \mathcal{K}_i} \hat{c}_{k_i} \psi_{k_i}(\dot{\boldsymbol{q}}_{aff,i}), \quad i = 1, ..., d \tag{6}$$

where $\dot{\boldsymbol{q}}_{aff,i} \in \mathbb{R}^{d_i}$ contains the velocities that affect the friction in joint i, and the parameters $\hat{c}_{k_i}$ are estimates of the wavelet transforms. The friction model can be regarded as d wavelet networks with one layer, each with d_i inputs, one output, nodes ψ_{k_i} and corresponding weights $\hat{c}_{k_i}$.

C. Lyapunov-based Adaptation Law Design

The WN friction model is linear in the parameters, which can readily be seen by rewriting (6) as

$$\hat{\boldsymbol{\tau}}_f(\dot{\boldsymbol{q}}) = \boldsymbol{\Psi}(\dot{\boldsymbol{q}})\hat{\boldsymbol{c}}_f$$

where

$$\boldsymbol{\Psi} = \begin{bmatrix} \{\psi_{k_1}\}_{k_1 \in \mathcal{K}_1} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \{\psi_{k_d}\}_{k_d \in \mathcal{K}_d} \end{bmatrix},$$

$$\hat{\boldsymbol{c}}_f = \begin{bmatrix} \{\hat{c}_{k_1}\}_{k_1 \in \mathcal{K}_1} & \cdots & \{\hat{c}_{k_d}\}_{k_d \in \mathcal{K}_d} \end{bmatrix}^T.$$

Let $\boldsymbol{c}_f^*$ denote the optimal value of $\hat{\boldsymbol{c}}_f$ for given index sets $\mathcal{K}_i$. An adaptation law must be found to drive $\tilde{\boldsymbol{c}}_f = \boldsymbol{c}_f^* - \hat{\boldsymbol{c}}_f$ towards zero.

Define $\boldsymbol{\epsilon} \in \mathbb{R}^d$ to be the modeling error, given by

$$\boldsymbol{\epsilon} = \boldsymbol{\tau}_f(\dot{\boldsymbol{q}}) - \boldsymbol{\Psi}(\dot{\boldsymbol{q}})\boldsymbol{c}_f^*.$$

The modeling error is bounded by $\epsilon_{max} = sup_{t \geq 0} ||\boldsymbol{\epsilon}(t)||$, which is a finite number provided that $\dot{\boldsymbol{q}}$ is bounded. To avoid parameter drift in the presence of $\boldsymbol{\epsilon}$ we require $||\hat{\boldsymbol{c}}_f|| \leq B$, where $B > 0$ is a constant. Furthermore, define $\boldsymbol{e} = \dot{\boldsymbol{q}} - \hat{\dot{\boldsymbol{q}}} \in \mathbb{R}^d$ to be the prediction error, where $\hat{\dot{\boldsymbol{q}}} \in \mathbb{R}^d$ is estimated angular velocity.

Proposition 1: In the presence of the modeling error $\boldsymbol{\epsilon}$, the series-parallel identification model

$$\ddot{\hat{\boldsymbol{q}}} = \boldsymbol{A}\dot{\hat{\boldsymbol{q}}} - \boldsymbol{A}\dot{\boldsymbol{q}} + \boldsymbol{M}^{-1}(\boldsymbol{\tau} - \boldsymbol{C}\dot{\boldsymbol{q}} - \boldsymbol{N} - \hat{\boldsymbol{\tau}}_f) \tag{7}$$

where $\boldsymbol{A} \in \mathbb{R}^{d \times d}$ is a designer-chosen Hurwitz matrix, and the parameter adaptation law

$$\dot{\hat{\boldsymbol{c}}}_f = -\dot{\tilde{\boldsymbol{c}}}_f = \begin{cases} -\boldsymbol{\Gamma}\boldsymbol{\Psi}^T\boldsymbol{e}, \text{ if } ||\hat{\boldsymbol{c}}_f|| \leq B \\ \text{or } ||\hat{\boldsymbol{c}}_f|| = B \text{ and } \boldsymbol{e}^T\boldsymbol{\Psi}\boldsymbol{\Gamma}\hat{\boldsymbol{c}}_f \geq 0 \\ -\boldsymbol{\Gamma}\boldsymbol{\Psi}^T\boldsymbol{e} + \frac{\boldsymbol{e}^T\boldsymbol{\Psi}\boldsymbol{\Gamma}\hat{\boldsymbol{c}}_f}{||\hat{\boldsymbol{c}}_f||^2}\hat{\boldsymbol{c}}_f, \\ \text{if } ||\hat{\boldsymbol{c}}_f|| = B \text{ and } \boldsymbol{e}^T\boldsymbol{\Psi}\boldsymbol{\Gamma}\hat{\boldsymbol{c}}_f < 0 \end{cases} \tag{8}$$

where $\boldsymbol{\Gamma} \in \mathbb{R}^{(K_1+...+K_d)\times(K_1+...+K_d)}$ is a designer-chosen positive definite diagonal matrix whose entries correspond to the adaptation gain of each parameter, ensure that $||\boldsymbol{e}||$ and $||\tilde{\boldsymbol{c}}_f||$ are bounded. If $\epsilon_{max} = 0$, $\lim_{t\to\infty} \boldsymbol{e}(t) = \boldsymbol{0}$ and $\lim_{t\to\infty} \dot{\tilde{\boldsymbol{c}}}_f(t) = \boldsymbol{0}$. If in addition $\boldsymbol{\Psi}$ is persistently exciting, then $\lim_{t\to\infty} \tilde{\boldsymbol{c}}_f(t) = \boldsymbol{0}$.

Partial proof: With (7) the prediction error dynamics become

$$\dot{\boldsymbol{e}} = \ddot{\boldsymbol{q}} - \ddot{\hat{\boldsymbol{q}}} = \boldsymbol{A}\boldsymbol{e} - \boldsymbol{M}^{-1}(\boldsymbol{\Psi}\tilde{\boldsymbol{c}}_f + \boldsymbol{\epsilon}). \tag{9}$$

Consider the candidate Lyapunov function

$$V(\boldsymbol{e}, \tilde{\boldsymbol{c}}_f) = \frac{1}{2}\boldsymbol{e}^T\boldsymbol{M}\boldsymbol{e} + \frac{1}{2}\tilde{\boldsymbol{c}}_f^T\boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{c}}_f.$$

Recall that $\dot{\boldsymbol{M}} \approx \boldsymbol{0}$ by Assumption 2. Let $\boldsymbol{Q} \in \mathbb{R}^{d\times d}$ be a positive definite matrix such that the Lyapunov equation $\boldsymbol{A}^T\boldsymbol{M} + \boldsymbol{M}\boldsymbol{A} = -\boldsymbol{Q}$ is satisfied. Differentiating V with respect to time, using (8) and (9), yields (after some calculation)

$$\dot{V} \leq -\frac{1}{2}\boldsymbol{e}^T\boldsymbol{Q}\boldsymbol{e} - \boldsymbol{e}^T\boldsymbol{\epsilon} \leq -\frac{1}{2}\boldsymbol{e}^T\boldsymbol{Q}\boldsymbol{e} + ||\boldsymbol{e}||\epsilon_{max}.$$

Hence if $\epsilon_{max} = 0$, $\dot{V}$ is negative semi-definite. If $\epsilon_{max} \neq 0$, then $\dot{V} > 0$ for $||\boldsymbol{e}|| < e_{min}$ where e_{min} is a constant that depends on $\boldsymbol{Q}$ and ϵ_{max}, thus implying the boundedness of $||\boldsymbol{e}||$. The proof of the boundedness of $||\tilde{\boldsymbol{c}}_f||$ and the validity of the limit properties are shown in [17] for an analogous problem and is omitted here. □

D. Dynamic Wavelet Network Structure

A strategy is adopted from [18] to automatically select the wavelets with the best dilations m_i and translations $\boldsymbol{n}_i$ as the robot travels within $\mathcal{X}_r$, in order to estimate the friction as accurately as possible using a minimum number of elements (wavelets) in (6). Before adaptation starts, a pool of wavelets is generated for each joint, corresponding to an allowed dilation range and wavelet centers inside $\mathcal{X}_r$. The pools are equivalent to the $\mathcal{K}_i$. During adaptation wavelets are considered for inclusion or exclusion based on a thresholding procedure. This procedure determines whether a wavelet ψ_{k_i} is allowed to be in the network or not, depending on the value of the corresponding parameter $\hat{c}_{k_i}(t)$ at time t.

Initially there are 2^{d_i} wavelets in the network[1]. The inclusion of wavelets is based on the assumption that adjacent wavelets in space and frequency are correlated. Thus, whenever the parameter $\hat{c}_{k_i}$ becomes greater than a threshold $\mu > 0$ in absolute value, all adjacent wavelets in $\mathbb{Z}^{1+d_i}$ are included in the network. Likewise, when the parameter $\hat{c}_{k_i}$ becomes smaller than a threshold $\theta > \mu$ in absolute value, and if it is still decreasing, the wavelet ψ_{k_i} is excluded from the network. The thresholds μ and θ are chosen by the designer and influence the accuracy of the estimation.

In the preceding Lyapunov analysis the network sizes were assumed to be constant. Inclusion of a new wavelet does not represent a problem since its parameter is initially set to 0. However, when excluding a wavelet its parameter suddenly jumps from a nonzero value to zero, which causes a discontinuity that may render the algorithm unstable. The problem is avoided by gradually setting the parameter to 0.

[1]For instance, if $\dot{\boldsymbol{q}}_{aff,i} \in \mathbb{R}^2$ a wavelet is placed in each quadrant of the 2-dimensional velocity plane.

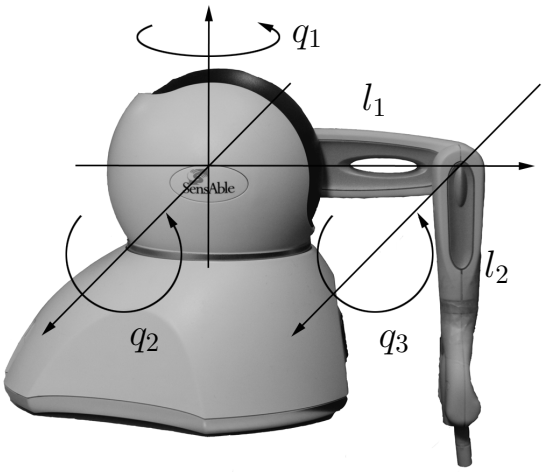

Fig. 2. Side view of the PHANTOM Omni robot.

IV. Experiments

The PHANTOM Omni haptic device from Sensable Technologies (see Fig. 2) was used to verify the WN friction model. The Omni is equipped with 6 angular position encoders, but the 3 wrist joints are not actuated, and so it is effectively a 3-DOF robot. Since the motor of joint 3 is located at the base, torque must be transmitted through cables inside link l_1 via a pulley at joint 2. The stiffness of the cables is assumed to be high enough for the position (and velocity) at the joint and base to be considered equal. However, this assumption may not hold whenever the robot is interacting with a stiff environment. Angular velocity $\dot{\boldsymbol{q}}$ has to be approximated from the joint angle $\boldsymbol{q}$, as only position encoder data are available. As the friction estimation theory was presented for continuous-time systems, all differential equations were integrated using Euler's method at 1 kHz for discrete-time implementation. As a performance metric the ability to compute the applied joint torque $\boldsymbol{\tau}$ in a reverse manner was used, i.e. according to

$$\hat{\boldsymbol{\tau}}_L = \boldsymbol{Y}_L(\dot{\boldsymbol{q}}, \boldsymbol{q})\boldsymbol{\alpha} + \hat{\boldsymbol{\tau}}_{f,L} \tag{10}$$

where the same low-pass filtering technique as in Section II is used to avoid acceleration measurements. Performance was quantified using relative root-mean-square (RMS) values, which for a signal $x(t)$ and its estimate $\hat{x}(t)$ is given by

$$RMS_{rel}(x, \hat{x}) = \frac{RMS(x - \hat{x})}{RMS(x)}.$$

A. Basic Dynamic Model

Building the dynamic model (2) for the Omni robot results in a $\boldsymbol{Y} \in \mathbb{R}^{3\times 9}$ regressor matrix and a $\boldsymbol{\alpha} \in \mathbb{R}^9$ parameter vector. To model the spring-like gravity compensation mechanism of the Omni a torsional spring $\tau_{spring} = \alpha_{10}(q_2 - \pi/2)$ is added to joint 2, where α_{10} is the spring constant. The final model is still linear in the parameters, with $\boldsymbol{Y} \in \mathbb{R}^{3\times 10}$ and $\boldsymbol{\alpha} \in \mathbb{R}^{10}$.

TABLE I

Steady-state $\boldsymbol{\alpha}$ values (mean value $\pm$ standard deviation).

$\boldsymbol{\alpha}$	Value	$\boldsymbol{\alpha}$	Value
α_1	$(1.178 \pm 0.034) \cdot 10^{-3}$	α_6	$(2.088 \pm 0.028) \cdot 10^{-3}$
α_2	$(6.395 \pm 0.240) \cdot 10^{-4}$	α_7	$(5.320 \pm 0.168) \cdot 10^{-4}$
α_3	$(-3.582 \pm 2.696) \cdot 10^{-5}$	α_8	$(2.043 \pm 0.078) \cdot 10^{-1}$
α_4	$(2.130 \pm 0.068) \cdot 10^{-3}$	α_9	$(7.015 \pm 0.025) \cdot 10^{-2}$
α_5	$(9.968 \pm 0.910) \cdot 10^{-5}$	α_{10}	$(1.493 \pm 0.068) \cdot 10^{-1}$

A least-squares method [15] was used to identify the elements of $\boldsymbol{\alpha}$. A PID controller was implemented for angular position control, with the 3 joint reference trajectories generated as sums of 4 sinusoids each for a total of 12 different frequencies, to ensure a high degree of excitation. The frequencies of the sinusoids were chosen to be relatively high[2], between 0.5Hz and 2Hz, for inertial effects to be dominant over friction. The low-pass filter cut-off frequency ω was set to 8Hz, in an attempt to place it between the system bandwidth and noise frequencies. Data were collected for 4 minutes, and the steady-state values of $\boldsymbol{\alpha}$ are shown in Table I. A basic CV model was used, but here only to absorb friction 'noise'.

B. Friction Estimation with Wavelet Networks

Only the friction in joint 1 of the Omni robot can be modeled individually as a function of $\dot{q}_1$. Because of the cable transmission friction in joints 2 and 3 is a function not only of $\dot{q}_2$ and $\dot{q}_3$, respectively, but also of the relative velocity $\dot{q}_{23} = \dot{q}_2 - \dot{q}_3$ between the joints. Hence, for the implementation of the WN friction model let $\dot{q}_{aff,1} = \dot{q}_1$ and $\dot{\boldsymbol{q}}_{aff,2} = \dot{\boldsymbol{q}}_{aff,3} = [\dot{q}_2, \dot{q}_3]^T$ in (6).

Three experiments were carried out in order to compare the performance of the CV friction model with the WN friction model. In experiments 1 and 2 the following protocol was adhered to. The data were collected by having the Omni robot follow sinusoidal trajectories, using the same PID controller as in the previous section, but for frequencies no higher than 0.7Hz. Each data series had a duration of 4 minutes. The data were analyzed off-line to estimate friction, using both the CV model and the proposed WN model, followed by the reverse torque computation with the adaptation turned off. Experiment 3 was for testing the trained WN friction model on an independent data set.

1) Experiment 1: In experiment 1 the effect of changing the relative velocity $\dot{q}_{23}$ between joints 2 and 3 on friction estimation was investigated. The purpose of the experiment was to show that the basic CV model

$$\hat{\tau}_{f,i}(\dot{q}_i) = \alpha_{c,i}\text{sgn}(\dot{q}_i) + \alpha_{v,i}\dot{q}_i, \quad i = 1, ..., 3,$$

which does not take the coupled friction into account, will fail when compared to the WN friction model. Two data series were collected, one in which the reference trajectories for joints 2 and 3 were equal in order to minimize $\dot{q}_{23}$, and one in which the reference trajectories were equal but

[2]Assumption 2 (low velocity) is only relevant for friction estimation.

TABLE II

RELATIVE RMS VALUES FOR EXPERIMENT 1.

RMS %	Small relative velocity			Large relative velocity		
	CV	WN	Diff.	CV	WN	Diff.
Joint 1	30.94	29.54	1.40	30.82	30.24	0.58
Joint 2	29.28	21.79	7.49	20.61	17.60	3.01
Joint 3	8.21	6.77	1.44	11.99	10.36	1.63

TABLE III

ESTIMATED COULOMB+VISCOUS PARAMETERS, EXPERIMENT 1.

Small relative velocity		Large relative velocity	
$\boldsymbol{\alpha}$	Value	$\boldsymbol{\alpha}$	Value
$\alpha_{c,1}$	$(1.980 \pm 0.015) \cdot 10^{-2}$	$\alpha_{c,1}$	$(2.004 \pm 0.013) \cdot 10^{-2}$
$\alpha_{v,1}$	$(1.849 \pm 0.189) \cdot 10^{-3}$	$\alpha_{v,1}$	$(1.552 \pm 0.167) \cdot 10^{-3}$
$\alpha_{c,2}$	$(2.011 \pm 0.042) \cdot 10^{-2}$	$\alpha_{c,2}$	$(2.381 \pm 0.321) \cdot 10^{-2}$
$\alpha_{v,2}$	$(5.323 \pm 0.548) \cdot 10^{-3}$	$\alpha_{v,2}$	$(1.187 \pm 0.034) \cdot 10^{-2}$
$\alpha_{c,3}$	$(1.689 \pm 0.017) \cdot 10^{-2}$	$\alpha_{c,3}$	$(2.003 \pm 0.020) \cdot 10^{-2}$
$\alpha_{v,3}$	$(5.987 \pm 0.172) \cdot 10^{-3}$	$\alpha_{v,3}$	$(1.097 \pm 0.032) \cdot 10^{-2}$

TABLE IV

WAVELET NETWORK THRESHOLDS AND DILATION RANGES.

	μ	θ	m
Joint 1	$4.0 \cdot 10^{-6}$	$5.0 \cdot 10^{-6}$	-5, ..., 0
Joint 2	$5.0 \cdot 10^{-5}$	$6.0 \cdot 10^{-5}$	-4, ..., 0
Joint 3	$5.0 \cdot 10^{-5}$	$6.0 \cdot 10^{-5}$	-4, ..., 0

with opposite signs in order to maximize $\dot{q}_{23}$. Each reference trajectory consisted of 4 different sinusoids.

The relative RMS values of the WN model were all lower than the CV model, as shown in Table II. Furthermore, note that when looking at the estimated parameters of the CV model in Table III, it can be seen that the Coulomb friction of joints 2 and 3 increased between the two data series. The basic CV model was fooled into perceiving greater friction when $\dot{q}_{23}$ was large. With the WN model, containing 2-dimensional wavelets, this was not a problem, since for the two data series we simply trained two different regions of the $\dot{q}_2$-$\dot{q}_3$ plane. These two regions corresponded to two straight lines through the origin at $\pm 45°$, representing equal—or equal with opposite signs—reference velocities.

The matrices $\boldsymbol{A}$ and $\boldsymbol{\Gamma}$, the thresholds μ and θ, and the allowed range of dilations m were chosen on the basis of trial and error. For instance, μ and θ have a direct influence on the size of the networks, which in these experiments could reach several hundred wavelets. The matrix $\boldsymbol{A}$ was chosen to give a fast convergence of the prediction error dynamics. The dilation range m determines how well the discontinuity around zero can be estimated. In experiment 1 $\boldsymbol{A}$ was set to $\boldsymbol{A} = -300 \cdot \boldsymbol{I}_{3\times 3}$ and the elements of $\boldsymbol{\Gamma}$ were set to be the same for all parameters in the same network, namely 0.08 for joint 1, 0.05 for joint 2 and 0.05 for joint 3. The values of μ, θ and m are shown in Table IV. For the second part of experiment 1, where the desired $\dot{q}_{23}$ was large, μ and θ had different, but similar values. In the experiments the constant B was only used as an indicator of instability (for example, caused by the choice of $\boldsymbol{A}$), and was set to a large number. No parameter drift was observed.

TABLE V

RELATIVE RMS VALUES FOR EXPERIMENT 2.

RMS %	CV	WN	Difference
Joint 1	30.18	29.76	0.42
Joint 2	21.00	20.50	0.50
Joint 3	12.97	9.02	3.95

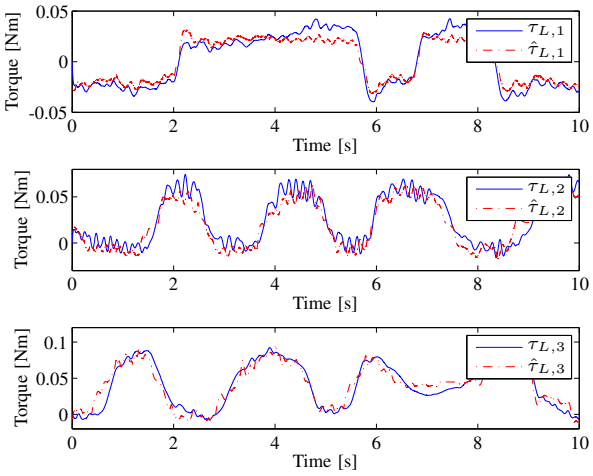

Fig. 3. Comparing low-pass filtered applied torque and its estimate.

2) Experiment 2: To show that a modified CV model that does take coupled friction into account will also fail against the WN friction model a second experiment was performed. For experiment 2 the CV model

$$\hat{\tau}_{f,2} = \alpha_{c,2}\text{sgn}(\dot{q}_2) + \alpha_{v,2}\dot{q}_2 + \alpha_{cr,2}\text{sgn}(\dot{q}_{23}) + \alpha_{vr,2}\dot{q}_{23}$$

was used for joint 2, and analogously for joint 3. Joint 1 was modeled as before. The training trajectories still consisted of 4 sinusoids for each joint, but in order to have a better training than in experiment 1 the sinusoids of joints 2 and 3 were independent. One series of data was obtained from the Omni, and the resulting RMS values are shown in Table V. As in the first experiment all RMS values of the WN model were lower than the CV model. The same values were used for $\boldsymbol{A}$, $\boldsymbol{\Gamma}$, m as in experiment 1, while similar values were used for μ and θ.

3) Experiment 3: In order to validate the model parameters found during training on an independent data series, the Omni was used as a slave robot in a teleoperation setup, and a random position trajectory was generated as the slave followed the master during manual manipulation. The reverse torque computation was performed on an extract of 10 seconds from the collected data series, using the model parameters found in experiment 2. A visual clue of the reverse torque computation performance of the WN model is provided in Fig. 3, showing the low-pass filtered applied torque and the estimated torque, as computed in (10).

Relative RMS values are shown in Table VI. The WN model still outperformed the CV model for joint 1. When testing 1-dimensional WNs one only needs to stay within

TABLE VI

RELATIVE RMS VALUES FOR EXPERIMENT 3.

RMS %	CV	WN	Difference
Joint 1	33.80	31.55	2.25
Joint 2	23.92	32.54	-8.62
Joint 3	16.59	20.10	-3.51

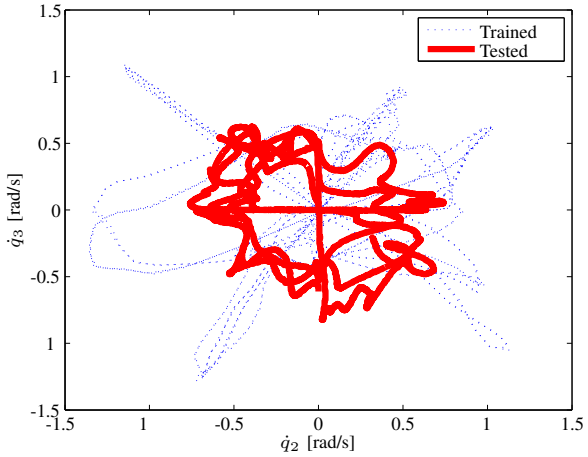

Fig. 4. Comparing trained and tested trajectories in the $\dot{q}_2$-$\dot{q}_3$ plane.

the trained velocity interval along the real line. However, for joints 2 and 3 the WN model performed worse than the CV model. These results indicate that the training of the 2-dimensional WNs in experiment 2 was not sufficient. Fig. 4 shows the velocity trajectory in the $\dot{q}_2$-$\dot{q}_3$ plane from training as a dotted line, while the trajectory from testing is shown as a thick, solid line. It is seen that the velocity during testing often found itself on untrained ground, especially at the bottom and to the left. It is emphasized that no optimal training sequence was sought in experiment 2, and in this way experiment 3 could serve as a good example of the unfortunate outcome of that.

V. CONCLUSION

A WN friction model has been proposed as a solution to the problem of finding a linear-in-parameter friction model for robots with coupled friction. The best fit wavelets were automatically selected with a thresholding procedure, and a Lyapunov approach was taken to design an adaptation law that drives the model parameters to the optimal values.

The performance of the WN model was tested and compared with the CV model through experiments with the PHANTOM Omni haptic device. During training the relative RMS values were found to be lower for the WN model than for the CV model in all cases, with an improvement ranging from 0.4 to 7.5 percentage points. At the same time insufficient training was also identified as one of the issues with our approach, implying that the price of avoiding friction modeling is a training challenge that requires careful attention.

The comparison of the proposed approach with other approaches with similar properties, such as neural or fuzzy techniques, will be considered in future works. The aim is to obtain a suitable friction model as part of the dynamic model to estimate contact forces when the robot is interacting with its environment.

ACKNOWLEDGEMENT

This work was partially funded by The Research Council of Norway project 167529/V30 and the E.U. contract MRTN-CT-2004-512400.

REFERENCES

[1] O. Egeland and J. T. Gravdahl, *Modeling and simulation for automatic control*. Trondheim, Norway: Tapir trykkeri, 2003.

[2] A. M. Tahmasebi, B. Taati, F. Mobasser, and K. Hashtrudi-Zaad, "Dynamic parameter identification and analysis of a PHANToM^{TM} haptic device," in *Proc. IEEE Conference on Control Applications (CCA'05)*, Toronto, Canada, Aug. 2005, pp. 1251–1256.

[3] P. R. Dahl, "A solid friction model," Aerospace Corporation, El Segundo, USA, Tech. Rep., May 1968.

[4] C. C. de Wit, H. Olsson, K. J. Åström, and P. Lischinsky, "A new model for control of systems with friction," *IEEE Trans. Automat. Contr.*, vol. 40, no. 3, pp. 419–425, Mar. 1995.

[5] J. Swevers, F. Al-Bender, C. G. Ganseman, and T. Prajogo, "An integrated friction model structure with improved presliding behavior for accurate friction compensation," *IEEE Trans. Automat. Contr.*, vol. 45, no. 4, pp. 675–686, Apr. 2000.

[6] P. Dupont, V. Hayward, B. Armstrong, and F. Altpeter, "Single state elastoplastic friction models," *IEEE Trans. Automat. Contr.*, vol. 47, no. 5, pp. 787–792, May 2002.

[7] G. M. Prisco and M. Bergamasco, "Dynamic modelling of a class of tendon driven manipulators," in *Proc. 8th IEEE International Conference on Advanced Robotics*, Monterey, USA, Jul. 1997, pp. 893–899.

[8] H. Kobayashi and R. Ozawa, "Adaptive neural network control of tendon-driven mechanisms with elastic tendons," *Automatica*, vol. 39, no. 9, pp. 1509–1519, Sep. 2003.

[9] C. C. de Wit, P. Noel, A. Aubin, B. Brogliato, and P. Drevet, "Adaptive friction compensation in robot manipulators: low velocities," in *Proc. IEEE International Conference on Robotics and Automation*, Scottsdale, USA, May 1989, pp. 1352–1357.

[10] F. Jatta, R. Adamini, A. Visioli, and G. Legnani, "Hybrid force/velocity robot contour tracking: an experimental analysis of friction compensation strategies," in *Proc. IEEE International Conference on Robotics and Automation*, Washington, DC, USA, May 2002, pp. 1723–1728.

[11] S. N. Huang, K. K. Tan, and T. H. Lee, "Adaptive friction compensation using neural network approximations," *IEEE Trans. Syst., Man, Cybern. C*, vol. 30, no. 4, pp. 551–557, Nov. 2000.

[12] H. Du and S. S. Nair, "Identification of friction at low velocities using wavelet basis function network," in *Proc. American Control Conference*, vol. 3, Philadelphia, USA, Jun. 1998, pp. 1918–1922.

[13] V. Santibañez, R. Kelly, and M. A. Llama, "A novel global asymptotic stable set-point fuzzy controller with bounded torques for robot manipulators," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 3, pp. 362–372, Jun. 2005.

[14] S. C. P. Gomes, D. da Silva Gomes, and C. M. Diniz, "Neuro-fuzzy friction compensation to robotic actuators," in *Proc. IEEE International Conference on Mechatronics*, Taipei, Taiwan, Jul. 2005, pp. 846–851.

[15] J.-J. E. Slotine and W. Li, *Applied nonlinear control*. Englewood Cliffs, USA: Prentice Hall, 1991.

[16] I. Daubechies, *Ten lectures on wavelets*. Pennsylvania, USA: Society for Industrial and Applied Mathematics - SIAM, 1992, CBMS-NSF Regional Conference Series in Applied Mathematics.

[17] M. M. Polycarpou and P. A. Ioannou, "Identification and control of nonlinear systems using neural network models: design and stability analysis," University of Southern California, Tech. Rep. 91-09-01, Sep. 1991.

[18] R. M. Sanner and J.-J. E. Slotine, "Structurally dynamic wavelet networks for the adaptive control of uncertain robotic systems," in *Proc.34th IEEE Conference on Decision and Control*, vol. 3, New Orleans, USA, Dec. 1995, pp. 2460–2467.