Asian Journal of Control, Vol. 16, No. 3, pp. 818–829, May 2014
Published online 26 June 2013 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/asjc.755

COMPARATIVE STUDY ON STATE ESTIMATION IN ELASTIC JOINTS

Wenjie Chen and Masayoshi Tomizuka

ABSTRACT

In robots with elastic joints, discrepancies exist between the motor side and the load side (*i.e.*, the link of the robotic joint) due to joint elasticity. Thus, sensing only at the motor side does not precisely provide the information of interest at the load side. In view that the load side states determine the position and orientation of the end-effector, it is critical to have the load side states sensed and/or estimated. In this paper, we present a comprehensive comparative study on load side state estimation utilizing two different models of robotic elastic joints. Low-cost MEMS inertia sensors, such as gyroscopes and accelerometers, are adopted on the load side to supplement the motor side sensing. We also incorporate sensor noises and biases in the measurement dynamics. Kalman filtering methods are designed based on the extended dynamic/kinematic models for the fusion of multiple sensor signals. We also study the specific issue related to the noise covariance adaptation. The effectiveness of the proposed schemes is experimentally demonstrated and compared on a robotic elastic joint testbed.

Key Words: Elastic joint, Kalman filtering, covariance adaptation.

I. INTRODUCTION

In control applications, discrepancies between the available measurements and the required information set difficulties in achieving good control performance. Such phenomena result from sensor dynamics as well as plant dynamics. Particularly, in robots with joint elasticity, due to gear compliances and nonlinearities, the ideal kinematic relationship (*e.g.*, the speed reduction ratio) between the load side (*i.e.*, the link of the robotic joint) and the motor side does not hold. This results in discrepancies between the actual physical states of the load side and the measurements at the motor side. Therefore, good motor side performance based on motor side measurements alone does not guarantee the desired good load side performance [1], which is a critical issue in practical applications.

The early work on the control of elastic joint robots dates back to the late 1980s. In [2], the dynamic model of elastic joint robots was formally introduced and the feedback linearization approach was proposed to control such robots with full state feedback including the load side states. Later it was claimed that a simple PD controller [3] or a set-point regulation controller [4], using motor side state feedback only, can globally stabilize robots with elastic joints. However, mismatches between the kinematics and dynamics of the actual robot and those of a model simplification deteriorate the performance of the motor side state feedback control due to the dependance on the dynamic model (*e.g.*, gravity torques and joint stiffness). Therefore, the load side state feedback is essential for good control performance, especially for accurate positioning of the end-effector and suppressing the vibration/resonance induced by joint compliances. Recently, several control approaches such as active disturbance rejection [5], multi-variable control loops [6], and linear pole placement [7], were proposed for vibration suppression, all of which confirmed the necessity of load side information feedback.

Precise load side position measurements (*e.g.*, load side encoder), however, are normally not available in industrial robots due to cost and assembly issues. Note that the load side resolution of the motor side encoder is improved by a factor of gear ratio. Thus, the load side encoder with the same load side resolution is normally much larger and much more expensive than the motor side encoder. To overcome this problem, inexpensive MEMS inertia sensors that are easy to mount may be considered on the load side instead of large and expensive load side encoders. A representative cost comparison between the encoders and the MEMS inertia sensors can be found in [8]. Consideration should be given, however, to problems such as non-negligible biases, limited bandwidth, and noises from low-cost sensors, which set restrictions on the direct utilization of sensor signals. These problems may be circumvented by the proper fusion of multiple sensor signals, to achieve clean and precise load side position estimation with substantial improvement over the motor side encoder measurements. Furthermore, the utilization of inertia sensors such as accelerometers and gyroscopes enables the

Manuscript received September 29, 2012; revised January 28, 2013; accepted March 31, 2013.

Wenjie Chen (corresponding author, e-mail: wjchen@me.berkeley.edu) and Masayoshi Tomizuka (e-mail: tomizuka@me.berkeley.edu) are with Department of Mechanical Engineering, University of California, Berkeley, CA 94720, USA.

This work was supported by FANUC Corporation, Japan. Real-time control hardware and software were provided by National Instruments, Inc.

direct measurement and/or estimation of the load side acceleration and velocity, which cannot be properly captured by a position encoder only.

The algorithms for load side state estimation have been extensively studied for robots with joint elasticity. In [9], a reduced order robust high gain observer was designed based on the robot dynamic model to estimate the load side position and velocity using motor side measurements only. In [10], a "semi-dynamic" model based observer was designed using motor encoder and load side accelerometer measurements. With the load side accelerometer measurement, the observer is decoupled into each joint and relies on only the motor and reducer related dynamic parameters. Recently, an extended-state observer was introduced in [11] to estimate the transformed coordinates (*i.e.*, load side position, velocity, acceleration, and jerk), in the presence of model uncertainty and external disturbance. However, the performance and stability of this observer may be poor when uncertainty/disturbance is fast varying. In [12], an extended Kalman filter (EKF) based on the dynamic model was proposed to estimate the load side states using motor side measurements. The complex robot dynamics and model uncertainty, however, greatly hinder the application of such EKF schemes to high-DOF (degree of freedom) robots.

The aforementioned methods are all based on the robot dynamic model requiring accurate system parameters, and thus are not reliable when subject to model uncertainties. To reduce the dependance on the dynamic model, some have approached a simpler and more accurate model representation, *i.e.*, the kinematic model. In [13,14], it was suggested that a Kalman filtering method based on the kinematic model (KKF) can be formulated by making use of accelerometers and position encoders. The model uncertainties are avoided by using the kinematic relation between the position and the acceleration. This method, however, was developed mainly for one-mass systems or direct drives (*i.e.*, without the joint compliance due to gear transmissions). Also, the main purpose was velocity estimation while assuming the position measurement is available.

Moreover, in most of the aforementioned methods, noise covariances were used as design parameters, which are not easy to select intuitively when multiple measurements are utilized. Non-negligible bias effect and limited bandwidth were also often ignored. Thus, it is desired to derive the estimation algorithms for robots with joint elasticity with further consideration of the physical sensor limitations.

In our previous work [15], we studied such estimation algorithms in two categories: dynamic model based and kinematic model based. Measurement dynamics were considered into the dynamic and kinematic models to deal with sensor noises and biases. The estimation algorithms were formulated based on these extended models to fuse the measurements from both the motor side and the load side. Adaptation of noise covariances was also addressed. In this paper, a more comprehensive comparative study for these estimation algorithms will be presented for elastic joints in the absence of precise load side position sensing. Complete formulation of the estimation schemes utilizing two different models (*i.e.*, dynamic and kinematic) will be given. A comparative experimental study is presented to validate the proposed schemes. The derivation of the covariance adaptation algorithm will be detailed. The schemes will also be discussed with potential extensions to accommodate non-negligible nonlinear dynamics.

Nomenclature

θ_ℓ	Load side position
$\theta_{\ell,s}$	Sensor measurement of load side position, θ_ℓ
$\dot{\theta}_{\ell,s}$	Sensor measurement of load side velocity, $\dot{\theta}_\ell$
$\ddot{\theta}_{\ell,s}$	Sensor measurement of load side acceleration, $\ddot{\theta}_\ell$
$\theta_{\ell f}$	Filtered load side position
$\theta_{\ell f,s}$	Fictitious measurement of filtered load side position, $\theta_{\ell f}$
θ_m	Motor side position
$\theta_{m,s}$	Sensor measurement of motor side position, θ_m
b_v	Bias of gyroscope measurement, $\dot{\theta}_{\ell,s}$
b_a	Bias of accelerometer measurement, $\ddot{\theta}_{\ell,s}$
u	Control input, *i.e.*, motor torque
u_s	Sensor measurement of motor torque, u
n_u	Measurement noise of motor torque, u
n_{sm}	Measurement noise of motor side position, θ_m
$n_{v\ell}$	Measurement noise of load side velocity, $\dot{\theta}_\ell$
$n_{a\ell}$	Measurement noise of load side acceleration, $\ddot{\theta}_\ell$
n_{ba}	Fictitious noise of accelerometer bias, b_a
n_{bv}	Fictitious noise of gyroscope bias, b_v
$n_{\ell f}$	Fictitious noise of filtered load side position (measurement), $\theta_{\ell f,s}$
d_j	Joint damping of the elastic joint
d_m	Motor side viscous damping coefficient
d_ℓ	Load side viscous damping coefficient
J_m	Motor side inertia of the elastic joint
J_ℓ	Load side inertia of the elastic joint
k_j	Joint stiffness of the elastic joint
N	Gear reduction ratio of the elastic joint

II. ESTIMATION ALGORITHMS

2.1 Model for measurement dynamics

The bias and the noise in the sensor output can be described as

$$\dot{b}(t) = n_b(t) \tag{1a}$$

$$y_s(t) = y(t) + b(t) + n_s(t) \tag{1b}$$

where b is the sensor bias varied by the fictitious noise, n_b. y is the actual physical quantity to be measured, y_s the sensor output, and n_s the measurement noise.

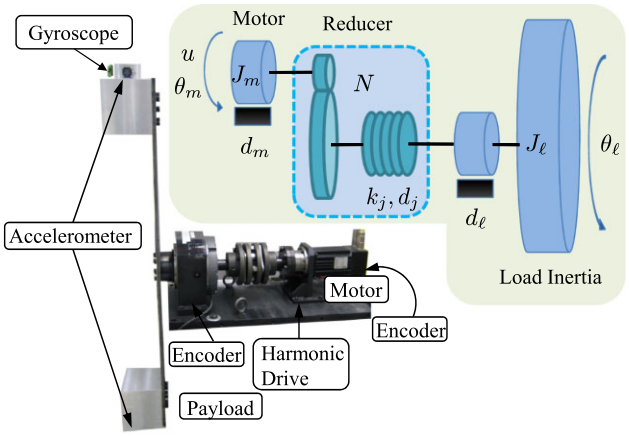

Fig. 1. A robotic elastic joint with gear transmission (the harmonic drive introduces a mismatch between the positions of the motor and the payload). The construction of this testbed will be detailed in Section 3.1.

2.2 Model for system dynamics

A typical robotic elastic joint system can be modeled as a two-mass system, as shown in Fig. 1. The subscripts m and ℓ denote the motor side and the load side quantities, respectively. d represents the viscous damping coefficient, and J is the moment of inertia. θ and u represent the angular position and the motor torque, respectively, and k_j and d_j are the stiffness and the damping coefficients of the gear reducer. The gear ratio of the reducer is denoted by N.

Equations of motion for the system in Fig. 1 are

$$J_m\ddot{\theta}_m + d_m\dot{\theta}_m = u - \frac{1}{N}\cdot\left[k_j\left(\frac{\theta_m}{N}-\theta_\ell\right)+d_j\left(\frac{\dot{\theta}_m}{N}-\dot{\theta}_\ell\right)\right] \tag{2a}$$

$$J_\ell\ddot{\theta}_\ell + d_\ell\dot{\theta}_\ell = k_j\left(\frac{\theta_m}{N}-\theta_\ell\right)+d_j\left(\frac{\dot{\theta}_m}{N}-\dot{\theta}_\ell\right) \tag{2b}$$

The state-space model representation of this elastic joint can be formulated as

$$\dot{x} = \mathbf{A}x + \mathbf{B}u. \tag{3}$$

where

$$x = [\theta_m \quad \dot{\theta}_m \quad \theta_\ell \quad \dot{\theta}_\ell]^{\mathrm{T}}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \dfrac{-k_j}{N^2J_m} & -\dfrac{d_m + d_j/N^2}{J_m} & \dfrac{k_j}{NJ_m} & \dfrac{d_j}{NJ_m} \\ 0 & 0 & 0 & 1 \\ \dfrac{k_j}{NJ_\ell} & \dfrac{d_j}{NJ_\ell} & \dfrac{-k_j}{J_\ell} & -\dfrac{d_j+d_\ell}{J_\ell} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 0 & \dfrac{1}{J_m} & 0 & 0 \end{bmatrix}^{\mathrm{T}}$$

The experimental setup of a robotic elastic joint in the Mechanical Systems Control Laboratory at the University of California, Berkeley, is shown in Fig. 1. As shown in the figure, the load side is equipped with both a gyroscope and a set of accelerometers. In this case, we introduce two state variables, b_a and b_v representing the biases of accelerometer and gyroscope outputs with the fictitious noises as n_{ba} and n_{bv}, respectively. Three measured signals are treated as the system outputs, *i.e.*, the motor side position ($\theta_{m,s}$), the load side velocity and acceleration ($\dot{\theta}_{\ell,s}$ and $\ddot{\theta}_{\ell,s}$), where $\bullet_s$ denotes the sensor measurement of the actual physical quantity $\bullet$. u in (3) represents the actual motor torque, which may differ from the commanded or measured torque u_s. It is assumed that u and u_s are related by $u_s = u + n_u$ (or $u = u_s - n_u$) where n_u is a fictitious noise.

After these considerations, the augmented dynamic model for the system in Fig. 1 becomes

$$\dot{x}_m(t) = \mathbf{A}_m x_m(t) + \mathbf{B}_{m,u}u_s(t) + \mathbf{B}_{m,w}w_m(t) \tag{4a}$$

$$y_m(t) = \mathbf{C}_m x_m(t) + v_m(t) \tag{4b}$$

where

$$x_m = \begin{bmatrix} x^{\mathrm{T}} & b_a & b_v \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{6\times 1}$$

$$y_m = \begin{bmatrix} \theta_{m,s} & \dot{\theta}_{\ell,s} & \ddot{\theta}_{\ell,s} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{3\times 1}$$

$$w_m = \begin{bmatrix} -n_u & n_{ba} & n_{bv} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{3\times 1}$$

$$v_m = \begin{bmatrix} n_{sm} & n_{v\ell} & n_{a\ell} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{3\times 1}$$

$$\mathbf{A}_m = \left[\begin{array}{c|c} A & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] \in \mathbb{R}^{6\times 6}$$

$$\mathbf{B}_{m,u} = \begin{bmatrix} B^{\mathrm{T}} & | & \mathbf{0} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{6\times 1}$$

$$\mathbf{B}_{m,w} = \left[\begin{array}{c|c} B & \mathbf{0} \\ \hline \mathbf{0} & I_2 \end{array}\right] \in \mathbb{R}^{6\times 3}$$

$$\mathbf{C}_m = \left[\begin{array}{cccc|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ \hline \multicolumn{4}{c|}{A_4} & 1 & 0 \end{array}\right] \in \mathbb{R}^{3\times 6}.$$

Here, A_4 is the fourth row of the matrix $\mathbf{A}$, and I_n is an $n \times n$ identity matrix. n_{sm}, $n_{v\ell}$, and $n_{a\ell}$ are the measurement

noises of motor side position, θ_m, load side velocity, $\dot{\theta}_\ell$, and load side acceleration, $\ddot{\theta}_\ell$, respectively.

The augmented model (4) naturally leads to a dynamic model-based Kalman filter with u_s as the input and y_m as the output. Instead of using a dynamic model, another way to construct a Kalman filter is to use a kinematic model, as detailed in the next section.

2.3 Model for system kinematics

The kinematic model from the acceleration to the position on the load side can be written as

$$\frac{d}{dt}\begin{bmatrix}\theta_\ell(t)\\ \dot{\theta}_\ell(t)\end{bmatrix}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\theta_\ell(t)\\ \dot{\theta}_\ell(t)\end{bmatrix}+\begin{bmatrix}0\\ 1\end{bmatrix}\ddot{\theta}_\ell(t) \tag{5}$$

The Kalman filter based on the kinematic model (5) is called the kinematic Kalman filter (KKF) [13]. In KKF, the acceleration measurement is used as an input to the filter, and the position measurement is used to correct the estimation output. In the robotic elastic joint, however, precise load side position measurement (*e.g.*, load side encoder) is normally not available, which makes it difficult to directly utilize the KKF algorithm.

Here, an original but intuitive method is proposed to approximate the load side position measurement. Let $G_{m2\ell}(s)$ be the transfer function from the motor side position θ_m to the load side position θ_ℓ, *i.e.*

$$G_{m2\ell}(s)\triangleq\frac{\theta_\ell(s)}{\theta_m(s)}=\frac{d_j s+k_j}{N\left[J_\ell s^2+(d_j+d_\ell)s+k_j\right]}$$

by the inherent system dynamics (2). Thus, $G_{m2\ell}(s)$ is approximately zero-phase static gain at low frequencies, since the dynamic chain from θ_m to θ_ℓ is modeled by mass, gear, spring, and damper. This indicates that the low frequency component of θ_ℓ can be approximated by that of θ_m.

Pass θ_m and θ_ℓ through a first order low pass filter $G_f(s)=\dfrac{\alpha}{s+\alpha}$, where $\alpha=\beta f_b$, f_b is the bandwidth of $G_{m2\ell}(s)$, and $\beta\in(0,\infty)$ is a design parameter for the low pass filter. Let the filter outputs be θ_{mf} and $\theta_{\ell f}$, respectively. It follows that

$$\dot{\theta}_{mf}=-\alpha\theta_{mf}+\alpha\theta_m \tag{6a}$$

$$\dot{\theta}_{\ell f}=-\alpha\theta_{\ell f}+\alpha\theta_\ell \tag{6b}$$

Therefore, if β is chosen properly, the filter outputs have the following relation

$$\theta_{\ell f}\approx\lambda_0\theta_{mf} \tag{7}$$

where $\lambda_0=\dfrac{1}{N}$ is the DC gain of $G_{m2\ell}(s)$. Generally, it is desired that $\beta\le 1$ and thus $\alpha\le f_b$. If the low frequency components of θ_m and θ_ℓ are highly dominant, however, β can be chosen to be larger than 1 to increase the filter convergence rate.

The filter dynamics (6) and the measurement dynamics (1) can be integrated into the system kinematic model (5), giving

$$\dot{x}_k(t)=\mathbf{A}_k x_k(t)+\mathbf{B}_{k,u}\ddot{\theta}_{\ell,s}(t)+\mathbf{B}_{k,w}w_k(t) \tag{8a}$$

$$y_k(t)=\mathbf{C}_k x_k(t)+v_k(t) \tag{8b}$$

where

$$x_k=\begin{bmatrix}\theta_{\ell f} & \theta_\ell & \dot{\theta}_\ell & -b_a & b_v\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{5\times1}$$

$$y_k=\begin{bmatrix}\theta_{\ell f,s} & \dot{\theta}_{\ell,s}\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{2\times1}$$

$$w_k=\begin{bmatrix}-n_{a\ell} & -n_{ba} & n_{bv}\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{3\times1}$$

$$v_k=\begin{bmatrix}n_{\ell f} & n_{v\ell}\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{2\times1}$$

$$\mathbf{A}_k=\left[\begin{array}{cccc|c}-\alpha & \alpha & 0 & 0 & \\ 0 & 0 & 1 & 0 & \mathbf{0}\\ 0 & 0 & 0 & 1 & \\ \hline \multicolumn{4}{c|}{\mathbf{0}} & \mathbf{0}\end{array}\right]\in\mathbb{R}^{5\times5}$$

$$\mathbf{B}_{k,u}=\begin{bmatrix}0 & 0 & 1 & 0 & 0\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{5\times1}$$

$$\mathbf{B}_{k,w}=\begin{bmatrix}\mathbf{0} \mid I_3\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{5\times3}$$

$$\mathbf{C}_k=\begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 1\end{bmatrix}\in\mathbb{R}^{2\times5}$$

and $n_{\ell f}$ is the fictitious noise of the filtered load side position (measurement), $\theta_{lf,s}$.

By the approximation in (7), $\lambda_0\theta_{mf,s}$, where $\theta_{mf,s}(s)=G_f(s)\theta_{m,s}(s)$, can be used as the fictitious measurement for the model output $\theta_{\ell f,s}$. The noise covariance of $\theta_{\ell f,s}$ can be approximated by that of $\lambda_0\theta_{mf,s}$, or used as a design parameter for the following Kalman filter.

2.4 Kalman filtering

The discrete time form of the extended system model (4) or (8) can be formulated as

$$x_d(k+1)=\mathbf{A}_d x_d(k)+\mathbf{B}_{d,u}u_d(k)+\mathbf{B}_{d,w}w_d(k) \tag{9a}$$

$$y_d(k)=\mathbf{C}_d x_d(k)+v_d(k) \tag{9b}$$

where $x_d(k)$, $y_d(k)$, $u_d(k)$, $w_d(k)$, and $v_d(k)$ are the k-th time step sampled values of $x_m(t)$, $y_m(t)$, $u_s(t)$, $w_m(t)$, and $v_m(t)$ for the

system dynamic model (4), or $x_k(t)$, $y_k(k)$, $\ddot{\theta}_{\ell,s}(t)$, $w_k(t)$, and $v_k(t)$ for the system kinematic model (8), respectively. $\mathbf{A}_d$, $\mathbf{B}_{d,u}$, $\mathbf{B}_{d,w}$, and $\mathbf{C}_d$ are derived from (4) or (8) by the zero-order-hold (ZOH) method [16]. In practice, $\mathbf{B}_{d,w}w_d$ could be generalized to include the disturbance and/or mismatched model dynamics.

For the extended system model (9), a Kalman filter to estimate the system states is given by

$$\hat{x}_d(k) = \hat{x}_d^o(k) + L_d(k)\tilde{y}_d^o(k) \tag{10a}$$

$$\hat{x}_d^o(k+1) = \mathbf{A}_d\hat{x}_d(k) + \mathbf{B}_{d,u}u_d(k) \tag{10b}$$

$$\tilde{y}_d^o(k) = y_d(k) - \mathbf{C}_d\hat{x}_d^o(k) \tag{10c}$$

where the Kalman filter gain L_d is calculated as

$$L_d(k) = M_d(k)\mathbf{C}_d^{\mathrm{T}}\left[\mathbf{C}_d M_d(k)\mathbf{C}_d^{\mathrm{T}} + \hat{R}_d(k)\right]^{-1} \tag{11a}$$

$$M_d(k+1) = \mathbf{A}_d Z_d(k)\mathbf{A}_d^{\mathrm{T}} + \mathbf{B}_{d,w}\hat{Q}_d(k)\mathbf{B}_{d,w}^{\mathrm{T}} \tag{11b}$$

$$Z_d(k) = M_d(k) - L_d(k)\mathbf{C}_d M_d(k) \tag{11c}$$

and $\hat{Q}_d(k)$ and $\hat{R}_d(k)$ are the covariance estimates of $w_d(k)$ and $v_d(k)$, respectively. $M_d(k)$ and $Z_d(k)$ are the one-step *a priori* and *a posteriori* covariances of the state estimation error, respectively. It can be checked that ($\mathbf{A}_d$, $\mathbf{C}_d$) is observable and ($\mathbf{A}_d$, $\mathbf{B}_{d,w}$) is controllable. The controllability by the disturbance input (e.g., w_d) rather than the control input (e.g., u_d) is important for the stationary property of Kalman filter. Note that ($\mathbf{A}_m$, $\mathbf{B}_{m,w}$) and ($\mathbf{A}_k$, $\mathbf{B}_{k,w}$) are controllable, and their ZOH equivalent ($\mathbf{A}_d$, $\mathbf{B}_{d,w}$) is also controllable when the sampling frequency (*i.e.*, 1 kHz) is specified. The observability of the discrete-time pair ($\mathbf{A}_d$, $\mathbf{C}_d$) can be checked similarly. Thus, if w_d and v_d are stationary zero-mean Gaussian white noises, L_d, M_d, and Z_d will converge to stationary matrices. However, w_d and v_d are often interpreted as fictitious noise terms and their covariances are adjusted to assign desired closed loop eigenvalues to the estimator [14]. In practice, they are chosen such that the estimator dynamics are five to ten times faster than that of the controller or fast enough to suppress the effect of perturbations [14]. Another way to deal with these covariances is to use the adaptive scheme described in the next section.

Remark 1. If only the accelerometer measurement is available on the load side for the estimation purpose, *i.e.*, no gyroscope measurement in the system output (4) or (8), the proposed method can still be formulated in a similar way. However, it is better to include the gyroscope measurement, since $\theta_{\ell f,s}$ uses a fictitious measurement $\lambda_0\theta_{mf,s}$ to formulate KKF, and $\dot{\theta}_{\ell,s}$ is the only load side real measurement in the KKF model output in (8). The benefits of this additional gyroscope will be demonstrated experimentally in Section III.

Remark 2. The KKF has several advantages [14] compared to the dynamic model based Kalman filter (DKF). First, the system representation using the kinematic model is simpler than the one using the dynamic model. Secondly, the kinematic model is an exact representation of the system states. It involves neither dynamic parameters nor external disturbances. Thus no model uncertainties need to be considered in the KKF, while the performance of the DKF method is subject to model uncertainties and nonlinear dynamics (*e.g.*, Coulomb friction and transmission error [17]) that are ignored when utilizing the linear dynamic model.

2.5 Estimation of noise covariance

As an optimal stochastic estimator, the Kalman filter assumes linear model and Gaussian additive noises with known covariances. In reality, however, it is often difficult to get the accurate covariances of the process or measurement noises. Thus, particular interest has been focused on the estimation of the noise covariances, *i.e.*, $\hat{Q}_d(k)$ and $\hat{R}_d(k)$.

In the proposed Kalman filter, the bias noises (n_{ba} and n_{bv}) and the output $\theta_{\ell f,s}$ are fictitious and their covariances cannot be physically identified. Practically, if the measurement bias is time invariant or slowly varying, it is desired that the fictitious bias covariance should be large enough at the beginning to steer the bias estimate quickly to its actual value. Then the fictitious noise covariance can be decreased to reduce the sensitivity of the bias estimation.

To handle uncertainties in these covariances, an adaptive approach utilizing the residual information for covariance estimation is applied. The routines presented here are modified from the algorithms summarized in [18,19]. The algorithm derivation using covariance matching technique is detailed in Appendix A.

By comparing the *a priori* and the *a posteriori* state estimates, the process noise covariance estimate $\hat{Q}_d(k)$ is updated by

$$\begin{aligned}\hat{Q}_d^*(k) = \mathbf{B}_{d,w}^{-1}[&\Delta x_d(k)\Delta x_d^{\mathrm{T}}(k) + Z_d(k) \\ &- \mathbf{A}_d Z_d(k-1)\mathbf{A}_d^{\mathrm{T}}](\mathbf{B}_{d,w}^{-1})^{\mathrm{T}}\end{aligned} \tag{12a}$$

$$\Delta x_d(k) = \hat{x}_d(k) - \hat{x}_d^o(k) \tag{12b}$$

$$\hat{Q}_d(k+1) = \hat{Q}_d(k) + [\hat{Q}_d^*(k) - \hat{Q}_d(k)]/N_Q \tag{12c}$$

where the adaptive estimate $\hat{Q}_d^*(k)$ is shown to be the residual pseudo-covariance plus the change between the *a priori* and the *a posteriori* covariances for the current time step. The next step covariance estimate $\hat{Q}_d(k+1)$ is then updated to this estimate $\hat{Q}_d^*(k)$ in an exponential moving average manner with N_Q as the window size. Note that, if $\mathbf{B}_{d,w}$ is not invertible, $\mathbf{B}_{d,w}^{-1}$ in (12a) can be replaced by the following pseudo-inverse

$$\mathbf{B}_{d,w}^{\dagger}=\begin{cases}\left(\mathbf{B}_{d,w}^{\mathrm{T}}\mathbf{B}_{d,w}\right)^{-1}\mathbf{B}_{d,w}^{\mathrm{T}}, & \mathbf{B}_{d,w}\text{ is thin;}\\ \mathbf{B}_{d,w}^{\mathrm{T}}\left(\mathbf{B}_{d,w}\mathbf{B}_{d,w}^{\mathrm{T}}\right)^{-1}, & \mathbf{B}_{d,w}\text{ is fat.}\end{cases} \tag{13}$$

Similarly, the measurement noise covariance estimate, $\hat{R}_d(k)$, can be adaptively updated by

$$\hat{R}_d^*(k)=\Delta y_d(k)\Delta y_d^{\mathrm{T}}(k)+\mathbf{C}_d Z_d(k)\mathbf{C}_d^{\mathrm{T}} \tag{14a}$$

$$\Delta y_d(k)=y_d(k)-\mathbf{C}_d\hat{x}_d(k) \tag{14b}$$

$$\hat{R}_d(k+1)=\hat{R}_d(k)+[\hat{R}_d^*(k)-\hat{R}_d(k)]/N_R \tag{14c}$$

where $\Delta y_d(k)$ is not the innovation but the *a posteriori* estimation error. The physical interpretation of this solution is that the theoretical value of measurement noise covariance should match with the estimation residual covariance (see Appendix A).

Remark 3. The above adaptation for the process (or measurement) noise covariance estimate, $\hat{Q}_d$ (or $\hat{R}_d$), is the sub-optimal solution developed separately assuming that the actual covariance R_d (or Q_d) is known [19] (see Appendix A). The two solutions can be used, with caution, to estimate both covariances simultaneously, or can be implemented one after another in the serial way, or can be confined to adapt $\hat{Q}_d$ (or $\hat{R}_d$) only.

Remark 4. The estimation performance is sensitive to the window sizes of the moving average filters. Thus, the window sizes should be carefully selected for each application. Sometimes it is preferred to use small window sizes at the beginning to speed up the covariance adaptation process. After the covariances have converged, the window sizes can be enlarged to maintain smooth estimation performance.

III. EXPERIMENTAL STUDY

3.1 Experimental setup

The proposed methods are implemented on a single-DOF elastic joint testbed shown in Fig. 1. This experimental setup consists of: (i) a servo motor with a 20,000 counts/revolution encoder; (ii) a harmonic drive with a 80:1 gear ratio; (iii) a load-side 144,000 counts/revolution encoder; and (iv) and a payload. The anti-resonant and resonant frequencies of the setup are approximately 11 Hz and 19 Hz, respectively.

The load side encoder is used only for performance evaluation. Besides the encoder, two other sensors are available on the load side. A MEMS gyroscope (Analog, ADXRS150) is installed at one end of the payload. Two accelerometers (Kistler, 8330A3) are installed at the two ends of the payload symmetrically as shown in Fig. 1 to compensate for the gravity effects on the measurements. Finally, the algorithms are implemented in a LabVIEW real-time target installed with LabVIEW Real-Time and FPGA modules. The sampling rate is selected as 1 kHz.

The noise of $\theta_{m,s}$ is bounded by the encoder resolution $\delta\theta_{m,s}$. This gives the approximate noise variance $\mathrm{var}(n_{sm})=\delta\theta_{m,s}^2/12$ [14]. The noise variances for $\ddot{\theta}_{\ell,s}$ and $\dot{\theta}_{\ell,s}$ can be obtained by zero-acceleration and zero-velocity experiments. The fictitious noise variances of $\theta_{\ell f,s}$, u_s, b_a, and b_v, *i.e.*, $\mathrm{var}(n_{\ell f,s})$, $\mathrm{var}(n_u)$, $\mathrm{var}(n_{ba})$, and $\mathrm{var}(n_{bv})$, cannot be determined experimentally (no torque measurement is available in this experimental setup), and thus can be used as design parameters or be determined with covariance adaptation method for the estimation performance.

3.2 Load side estimation experiments

The performance validation is conducted for the following five estimation configurations:

- *DKF-AG*: Dynamic model (4) based Kalman filter with accelerometers and gyroscope
- *DKF-A*: Dynamic model (4) based Kalman filter with accelerometers only
- *KKF-AG*: Kinematic model (8) based Kalman filter with accelerometers and gyroscope
- *KKF-A*: Kinematic model (8) based Kalman filter with accelerometers only
- *Torsion*: Use motor encoder output $\theta_{m,s}$ as the load side position directly. In this case the load side position estimation error is given by the joint torsion $\delta_{\ell,s}=\dfrac{\theta_{m,s}}{N}-\theta_{\ell,s}$.

3.2.1 Open loop frequency rich test without covariance adaption or parametric uncertainty

The effectiveness of the above algorithm configurations in the open loop frequency rich response is tested first. The motor torque command is a quadratic chirp signal with the magnitude of 0.2 Nm. The frequency range varies from 0.5 Hz to 50 Hz quadratically in 50 sec as shown in Fig. 2.

In this experiment, the algorithms are tested without covariance adaptation or parametric uncertainty to see the general benefits of load side sensors. The fixed noise variances and the dynamic parameter values used in the algorithms are listed in Table I. These variances and parameter values are mostly the nominal values identified experimentally as described in Section 3.1, except for those of $n_{\ell f,s}$, n_u, n_{ba}, and n_{bv}, which are tuned by trial and error. β in the low-pass filter in KKF algorithms is selected as 1 (*i.e.*, $\alpha=f_b$).

The estimation errors of the load side position by five estimation configurations are shown in Fig. 3. Note that the errors are of the same increasing frequency as the chirp torque input. In Fig. 3, the rms value on the right top of each

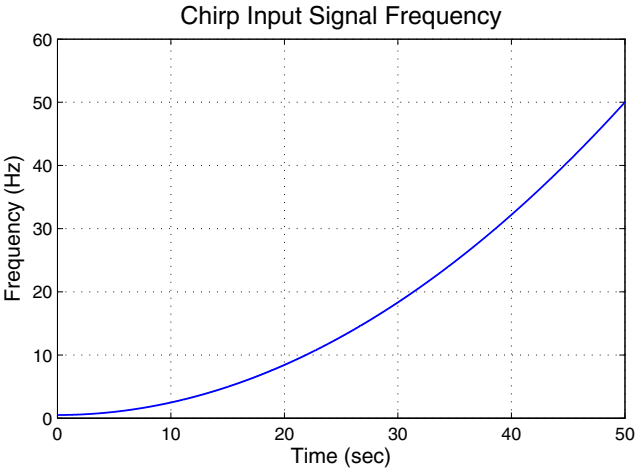

Fig. 2. Chirp input signal frequency.

Table I. Nominal parameter values—first experiment (SI units).

var(n_u)	var(n_{sm})	var($n_{\ell f,s}$)
10^{-4}	8.225×10^{-9}	8.225×10^{-9}
var($n_{v\ell}$)	var($n_{a\ell}$)	var(n_{bv})
2.525×10^{-5}	3.284×10^{-3}	10^{-5}
var(n_{ba})	J_m	J_ℓ
10^{-2}	5.313×10^{-4}	6.8
k_j	d_j	
3.1×10^4	47	

sub-figure denotes the root-mean-square value of the estimation error. As expected, the *KKF-AG* method achieves the best estimation performance, the rms estimation error of which is less than $\frac{1}{3}$ of the one by directly using θ_m (*i.e.*, Torsion). The utilization of the gyroscope does not benefit the dynamic model based algorithms, but results in significant improvement for the kinematic model based algorithms. This makes sense since the dynamic model (4) uses $\theta_{m,s}$, $\dot{\theta}_{\ell,s}$, and $\ddot{\theta}_{\ell,s}$ as the model output, all of which are available actual measurements, while the kinematic model (8) uses only $\theta_{\ell f,s}$ and $\dot{\theta}_{\ell,s}$ as the model output, and $\theta_{\ell f,s}$ is actually approximated by $\frac{\theta_{mf,s}}{N}$. With this benefit of the gyroscope in mind, in the following testing, performance comparisons are conducted only for configurations where both accelerometers and gyroscope are available, *i.e.*, *DKF-AG* and *KKF-AG*.

3.2.2 Open loop frequency rich test with covariance adaption and parametric uncertainty

In this experiment, the same open loop frequency rich response is utilized. The covariance estimation algorithm is

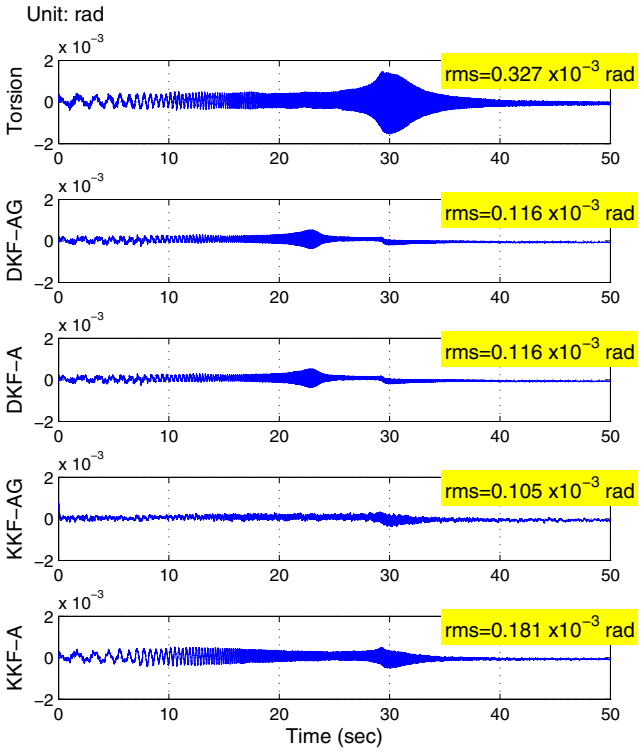

Fig. 3. Estimation errors of load side position in chirp experiment (without covariance adaptation or parametric uncertainty).

tested with the same initial noise variances listed in Table I except that the variances of n_{bv} and n_{ba} are reset to 10^{-4} and 0.1 (different from the tuned suboptimal values). Also, J_ℓ and k_j are varied with 20% error to check the performance of DKF when subject to the parametric uncertainty. β in the KKF method is selected as 0.5, *i.e.*, $\alpha = 0.5 f_b$. The adaptive covariance estimation scheme is confined to adapt $\hat{R}_d$ only.

Fig. 4 shows that *KKF-AG* still achieves the best estimation performance, where the estimation error is greatly reduced at all the testing frequencies. *DKF-AG* does not perform as well as *KKF-AG*, and it amplifies the estimation error at some frequency range due to parametric uncertainties.

The adapted $\hat{R}_d$ in both methods is shown in Fig. 5. As expected, to match the covariance with the estimation results, $\hat{R}_d$ tends to increase when the estimation error is increasing (especially around the resonant frequency 19 Hz at about 30 sec).

Fig. 6 shows that both *KKF-AG* and *DKF-AG* are effective to estimate the biases of gyroscope and accelerometer outputs. Thus, the bias effects in the utilization of sensor signals are attenuated.

In Fig. 6, the small fluctuation effect at the end of the trajectory for gyroscope bias estimation is mainly

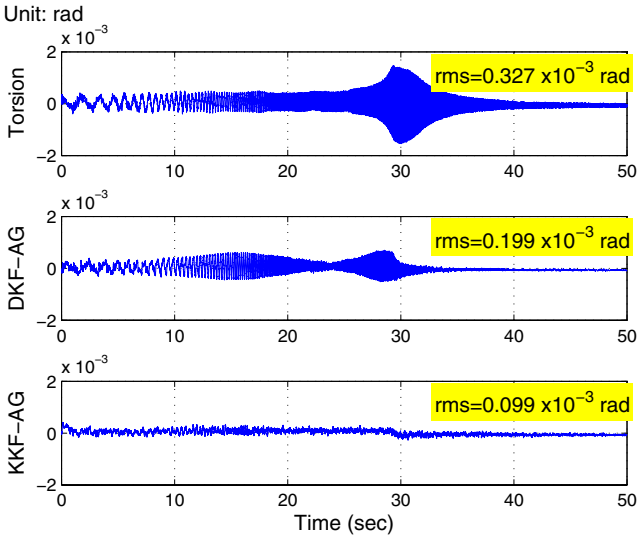

Fig. 4. Estimation errors of load side position in chirp experiment (with covariance adaptation and parametric uncertainty).

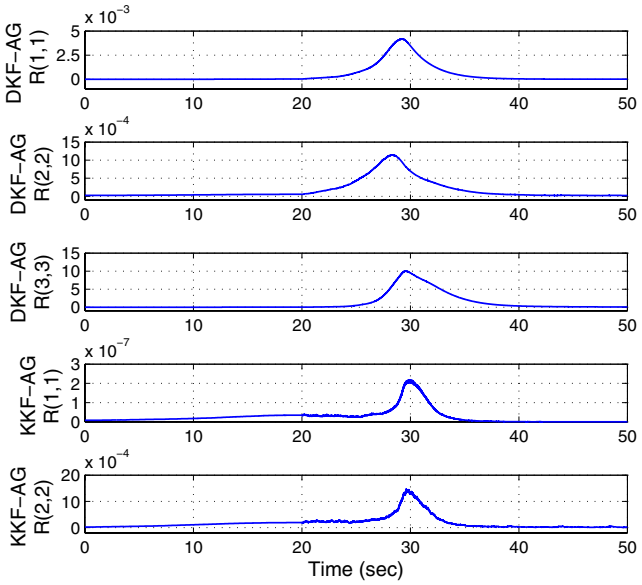

Fig. 5. Covariance estimation in chirp experiment (with covariance adaptation and parametric uncertainty).

due to the limited bandwidths of the sensors. The 3 dB bandwidth is 40 Hz for the gyroscope (Analog, ADXRS150) and 2700 Hz for the accelerometer (Kistler, 8330A3). When the chirp frequency increases to around 40 Hz (see Fig. 2) towards the end of the trajectory, the gyroscope measurement becomes not as accurate as desired. This results in small fluctuation for the gyroscope bias

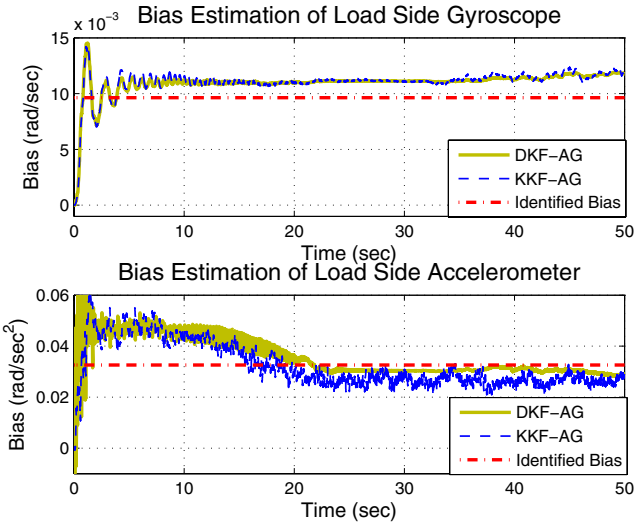

Fig. 6. Bias estimation in chirp experiment (with covariance adaptation and parametric uncertainty).

estimate in Fig. 6, while the accelerometer bias estimation continues similar performance as in the low frequency region.

Thanks to the adaptation algorithm for measurement noise covariances, this fluctuation behavior does not really affect the load side position estimation accuracy (see Fig. 4). Also note that, in Fig. 6, the accelerometer bias estimate by *KKF-AG* is not as good as *DKF-AG* at the end part. This is because that *KKF-AG* treats the accelerometer measurement as the control input in the model rather than as the measurement output (see (8)). In this paper, we confined the covariance adaptation to measurement noises only. Thus, the accelerometer noise covariance is not adapted in *KKF-AG* while it is in *DKF-AG*. This actually again validates the necessity of covariance adaptation in the utilization of such sensors with limited bandwidth.

On the other hand, the motions in actual robotic tasks are normally at a frequency region much lower than 40 Hz. Therefore, this bias estimation fluctuation will not be a problem for actual robotic tasks.

3.2.3 Scanning trajectory test with covariance adaption and parametric uncertainty

This experiment is to verify the effectiveness of the methods during a scanning trajectory. The desired load side trajectory (Fig. 7) is designed as a fourth-order smooth time optimal trajectory suggested in [20]. The controller for trajectory tracking is a basic PID feedback controller with a model based feedforward controller. All the covariances and dynamic parameter values remain the same as those in the second chirp experiment (Fig. 4) except for the window size

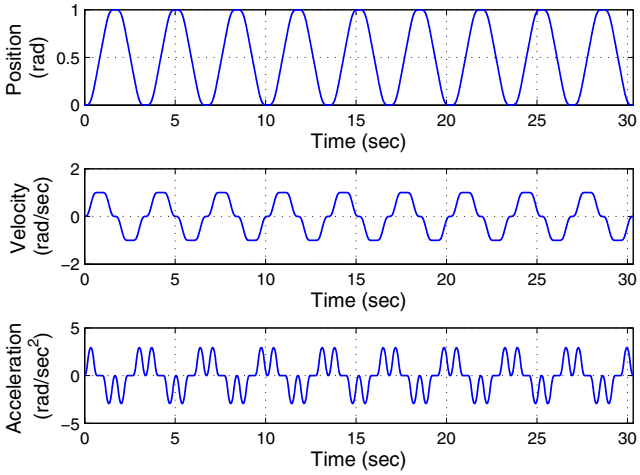

Fig. 7. Load side desired trajectory.

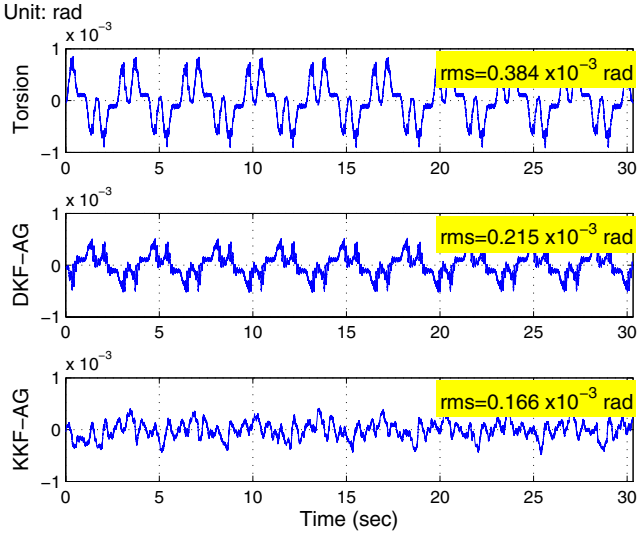

Fig. 8. Estimation errors of load side position in scanning trajectory experiment.

N_R used in the covariance adaptation scheme. This experiment is to test the load side estimation only and thus the estimated load side information is not fed back for control. This controller utilizes the motor side encoder feedback and is to achieve the motor side tracking, *i.e.*, $\theta_m \to \theta_{md}$, where θ_{md} is computed from the desired load side trajectory $\theta_{\ell d}$ (see [1] for controller details as well as the utilization of load side estimates for feedback control.

The result of this experiment is shown in Fig. 8, where *KKF-AG* still achieves the best estimation performance. It also indicates that the cleanness of the load side position estimate by *KKF-AG* is similar to the one by *Torsion*, which is computed by the encoder measurements and thus is clean

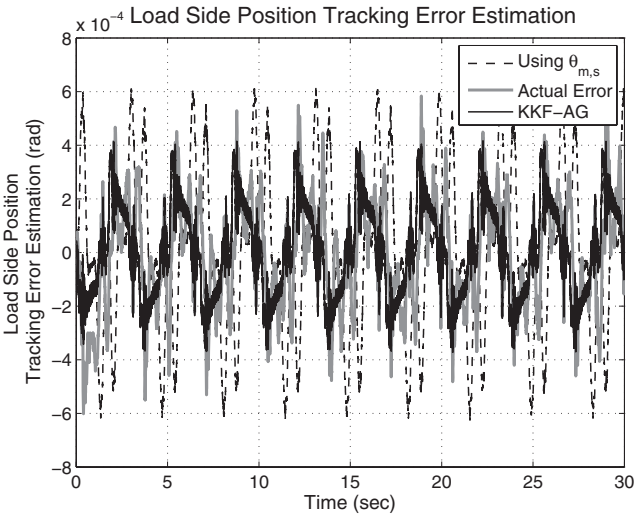

Fig. 9. Load side actual and estimated position tracking error.

at the encoder resolution level. Also note that the remaining oscillatory estimation error by *KKF-AG* is mainly due to the transmission error [17], which is not considered in the estimation algorithm. Fig. 9 illustrates the actual and the estimated tracking errors of the load side position. It shows that the tracking error estimated by *KKF-AG* captures most trends of the actual tracking error. This indicates the potential of *KKF-AG* that could benefit the load side position tracking application. To demonstrate this, the load side position estimate by *KKF-AG* has been utilized for real-time feedback control in the friction compensation in [1] and [15].

IV. DISCUSSION

The above experimental study has demonstrated the advantages of KKF over DKF. It is noted that the DKF and KKF methods are individually formulated based on different models. However, it may be beneficial to further combine both the dynamic model and the kinematic model for Kalman filter formulation. The following two ways may be used to achieve this objective.

First, in KKF, instead of simple first order low pass filter $G_f(s) = \dfrac{\alpha}{s+\alpha}$, the nominal dynamic model

$$\left[\frac{\theta_\ell(s)}{\theta_m(s)}\right]_{\text{nom}} = \frac{\hat{d}_j s + \hat{k}_j}{N\left[\hat{J}_\ell s^2 + (\hat{d}_j + \hat{d}_\ell)s + \hat{k}_j\right]} \tag{15}$$

can be used to obtain the load side position approximation from the motor side encoder measurement, where $\hat{\bullet}$ denotes the nominal value of $\bullet$. However, this will result in a higher

order model formulation which introduces more complexity and computation load. Actually, it may not be favorable to use a high order model since its high frequency response is always subject to large errors due to model uncertainties. Furthermore, one can note that nonlinear dynamic terms such as Coulomb friction and transmission error are still not considered in this method due to the utilization of the linear transfer function.

The second way to approximate the load side position measurement is to use the dynamic model (2) directly, which yields

$$\theta_{\ell,s} = \frac{1}{\hat{d}_j s + \hat{k}_j}\left[\frac{\hat{k}_j}{N}\theta_{m,s} + \frac{\hat{d}_j}{N}\dot{\theta}_{m,s} - N\left(u_s + \hat{f}_m - \hat{J}_m\hat{\ddot{\theta}}_{m,s} - \hat{d}_m\dot{\theta}_{m,s}\right)\right] \tag{16}$$

where $\theta_{m,s}$ and $\dot{\theta}_{m,s}$ are obtained from motor encoder measurements. u_s can be either motor torque command or measured by motor current. The desired trajectory $\ddot{\theta}_{md}$ can be used instead of $\hat{\ddot{\theta}}_{m,s}$ in (16) as approximation. By this approach, the nonlinear dynamic terms (*e.g.*, Coulomb friction and transmission error) can be taken into account as a lumped term $\hat{f}_m$ if a good model knowledge is available. This first order model formulation is relatively easy and efficient to realize. This extension has been adopted in [21] for the load side state estimation in the multi-joint robot.

V. CONCLUSIONS

This paper investigated the estimation algorithms for load side information of robotic elastic joints with a multi-sensor configuration. Kalman filtering methods were formulated based on the extended dynamic/kinematic models taking measurement dynamics (bias and noise) into consideration. Noise covariance adaptation was also studied. The effectiveness of the methods was verified experimentally on a single-DOF elastic joint testbed. The benefits of load side sensing and the advantages of KKF over DKF have been demonstrated.

VI. APPENDIX A

6.1 COVARIANCE ADAPTATION

Consider the discrete time model (9) with the assumption that the initial condition $x_d(1) \sim X_d(1) = \mathcal{N}\left(\hat{x}_d^o(1), M_d(1)\right)$, the process noise $w_d(k) \sim W_d(k) = \mathcal{N}(0, Q_d(k))$, and the measurement noise $v_d(k) \sim V_d(k) = \mathcal{N}(0, R_d(k))$. The Kalman filter to estimate the system states is given by (10)–(11). If noise covariances $Q_d(k)$ and $R_d(k)$ are unknown, the real-time estimation ((12) and (14)) for $Q_d(k)$ and $R_d(k)$ can be derived based on the maximum likelihood principle [19,18]. Furthermore, if noise covariances Q_d and R_d are time invariant constants, expectation maximization method [22,23] can be well applied to iteratively estimate these covariances in an off-line manner. Here, however, since the covariances of fictitious noises (Q_d and R_d) are likely to be time varying depending on the trajectory, we want to estimate them online. Therefore, only sub-optimal maximum likelihood solutions can be derived to estimate Q_d and R_d separately. There are other adaptive filtering methods introduced in [19,18], such as Bayesian methods, correlation methods, and covariance matching. Among these methods, covariance matching typically requires the least computation and storage although the convergence property may be subject to question in certain cases. It is, however, worth pointing out that the same sub-optimal maximum likelihood solutions for this problem can be derived by the covariance matching technique as detailed below.

Note that, if $Q_d(k)$ and $R_d(k)$ are known constants, the Kalman filter yields the optimal estimation for this problem. It follows that

$$\begin{aligned} E\left[\Delta x_d(k)\Delta x_d^{\mathrm{T}}(k)\right] &= E\left[L_d(k)\tilde{y}_d^o(k)\tilde{y}_d^o(k)^{\mathrm{T}}L_d^{\mathrm{T}}(k)\right] \\ &= L_d(k)\left(\mathbf{C}_d M_d(k)\mathbf{C}_d^{\mathrm{T}} + R_d(k)\right)L_d^{\mathrm{T}}(k) \\ &= L_d(k)\mathbf{C}_d M_d(k)\mathbf{C}_d^{\mathrm{T}}L_d^{\mathrm{T}}(k) + Z_d(k)\mathbf{C}_d^{\mathrm{T}}L_d^{\mathrm{T}}(k) \\ &\quad \left(\text{since } L_d(k) = Z_d(k)\mathbf{C}_d^{\mathrm{T}}R_d^{-1}(k)\right) \\ &= L_d(k)\mathbf{C}_d M_d(k)\mathbf{C}_d^{\mathrm{T}}L_d^{\mathrm{T}}(k) \\ &\quad + \left(I - L_d(k)\mathbf{C}_d\right)M_d(k)\mathbf{C}_d^{\mathrm{T}}L_d^{\mathrm{T}}(k) \\ &= M_d(k)\mathbf{C}_d^{\mathrm{T}}L_d^{\mathrm{T}}(k) = L_d(k)\mathbf{C}_d M_d(k) \end{aligned} \tag{17}$$

Hereafter the conditional expectation $E[\bullet|j]$ given the information up to the j-th time step is denoted as $E[\bullet]$ for simplicity.

Thus,

$$\begin{aligned} Z_d(k) &= M_d(k) - L_d(k)\mathbf{C}_d M_d(k) \\ &= M_d(k) - E\left[\Delta x_d(k)\Delta x_d^{\mathrm{T}}(k)\right] \\ \Rightarrow M_d(k) &= E\left[\Delta x_d(k)\Delta x_d^{\mathrm{T}}(k)\right] + Z_d(k) \\ \text{Also } M_d(k) &= \mathbf{A}_d Z_d(k-1)\mathbf{A}_d^{\mathrm{T}} + \mathbf{B}_{d,w}Q_d(k)\mathbf{B}_{d,w}^{\mathrm{T}} \\ \Rightarrow \mathbf{B}_{d,w}Q_d(k)\mathbf{B}_{d,w}^{\mathrm{T}} &= E\left[\Delta x_d(k)\Delta x_d^{\mathrm{T}}(k)\right] + Z_d(k) \\ &\quad - \mathbf{A}_d Z_d(k-1)\mathbf{A}_d^{\mathrm{T}} \end{aligned} \tag{18}$$

This means that the theoretical value for $Q_d(k)$ should match the actual estimation residual covariance $E\left[\Delta x_d(k)\Delta x_d^{\mathrm{T}}(k)\right]$ as in (18).

Similarly, the theoretical matching value for $R_d(k)$ can be obtained. Note that,

$$
\begin{aligned}
E\left[\Delta y_d(k)\,\Delta y_d^{\mathrm{T}}(k)\right] &= E\Big[\left(\mathbf{C}_d\left(x_d(k)-\hat{x}_d(k)\right)+v_d(k)\right)\\
&\quad\cdot\left(\mathbf{C}_d\left(x_d(k)-\hat{x}_d(k)\right)+v_d(k)\right)^{\mathrm{T}}\Big]\\
&= E\Big[\mathbf{C}_d\left(x_d(k)-\hat{x}_d(k)\right)\left(x_d(k)-\hat{x}_d(k)\right)^{\mathrm{T}}\mathbf{C}_d^{\mathrm{T}}\\
&\quad+\mathbf{C}_d\left(x_d(k)-\hat{x}_d(k)\right)v_d^{\mathrm{T}}(k)\\
&\quad+v_d(k)\left(x_d(k)-\hat{x}_d(k)\right)^{\mathrm{T}}\mathbf{C}_d^{\mathrm{T}}+v_d(k)\,v_d^{\mathrm{T}}(k)\Big]\\
&= \mathbf{C}_d Z_d(k)\mathbf{C}_d^{T}+E\left[\mathbf{C}_d\left(x_d(k)-\hat{x}_d(k)\right)v_d^{\mathrm{T}}(k)\right]\\
&\quad+E\left[v_d(k)\left(x_d(k)-\hat{x}_d(k)\right)^{\mathrm{T}}\mathbf{C}_d^{\mathrm{T}}\right]+R_d(k)
\end{aligned}
\tag{19}
$$

Furthermore,

$$
\begin{aligned}
&E\left[\mathbf{C}_d\left(x_d(k)-\hat{x}_d(k)\right)v_d^{\mathrm{T}}(k)\right]\\
&\quad= E\left[\mathbf{C}_d\left(x_d(k)-\hat{x}_d^{o}(k)-L_d(k)\left(y_d(k)-\mathbf{C}_d x_d^{o}(k)\right)\right)v_d^{\mathrm{T}}(k)\right]\\
&\quad= E\left[\mathbf{C}_d\left(\left(I-L_d(k)\mathbf{C}_d\right)\left(x_d(k)-\hat{x}_d^{o}(k)\right)-L_d(k)\,v_d(k)\right)v_d^{\mathrm{T}}(k)\right]\\
&\quad= E\left[-\mathbf{C}_d L_d(k)\,v_d(k)\,v_d^{\mathrm{T}}(k)\right]=-\mathbf{C}_d L_d(k)\,R_d(k)\\
&\quad= -\mathbf{C}_d Z_d(k)\mathbf{C}_d^{\mathrm{T}}\quad\left(\text{since } L_d(k)=Z_d(k)\mathbf{C}_d^{\mathrm{T}}R_d^{-1}(k)\right)
\end{aligned}
$$

Therefore, (19) is further reduced to

$$
\begin{aligned}
E\left[\Delta y_d(k)\Delta y_d^{\mathrm{T}}(k)\right] &= \mathbf{C}_d Z_d(k)\mathbf{C}_d^{\mathrm{T}}-2\mathbf{C}_d Z_d(k)\mathbf{C}_d^{\mathrm{T}}+R_d(k)\\
\Rightarrow R_d(k) &= E\left[\Delta y_d(k)\Delta y_d^{\mathrm{T}}(k)\right]+\mathbf{C}_d Z_d(k)\mathbf{C}_d^{\mathrm{T}}
\end{aligned}
\tag{20}
$$

For estimation based on the above covariance matching analysis, the expectations in (18) and (20) can be approximated by the average over the most recent N_s samples. An *ad hoc* way to reduce the filter memory requirement is by replacing the expectation with the most recent data only (i.e., $N_s = 1$) and then implementing the update in an exponential moving average form. This leads to the solution in (12) and (14).

REFERENCES

1. Chen, W., K. Kong, and M. Tomizuka, "Hybrid Adaptive Friction Compensation of Indirect Drive Trains," *Proc. 2009 ASME Dyn. Syst. Control Conf., Hollywood, California*, pp. 313–320 (2009).
2. Spong, M. W., "Modeling and control of elastic joint robots," *J. Dyn. Syst. Meas. Control Trans. ASME*, Vol. 109, No. 4, pp. 310–318 (1987).
3. Tomei, P., "A simple pd controller for robots with elastic joints," *IEEE Trans. Autom. Control*, Vol. 36, pp. 1208–1213 (1991).
4. Ailon, A. and R. Ortega, "An observer-based set-point controller for robot manipulators with flexible joints," *Syst. Control Lett.*, Vol. 21, No. 4, pp. 329–335, (1993).
5. Zhao, S. and Z. Gao, "An active disturbance rejection based approach to vibration suppression in two-inertia systems," *Asian J. Control*, Vol. 15, No. 2, pp. 350–362 (2013).
6. Feliu, V., F. Castillo, V. Jaramillo, and G. Partida, "A robust controller for a 3-dof flexible robot with a time variant payload," *Asian J. Control*, Vol. 15, No. 4, pp. 1–12 (2013).
7. Hu, J.-S., F.-R. Hu, and C.-H. Kang, "On the two-inertia system: Analysis of the asymptotic behaviors to multiple feedback position control," *Asian J. Control*, Vol. 15, No. 6, pp. 1–13 (2013).
8. Roan, P., N. Deshpande, Y. Wang, and B. Pitzer, "Manipulator state estimation with low cost accelerometers and gyroscopes," In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 4822–4827 (2012).
9. Jankovic, M., "Observer based control for elastic joint robots," *IEEE Trans. Robot. Autom.*, Vol. 11, pp. 618–623 (1995).
10. De Luca, A., D. Schroder, and M. Thummel, "An acceleration-based state observer for robot manipulators with elastic joints," *Robot. Autom. 2007 IEEE Int. Conf., Rome, Italy*, pp. 3817–3823 (2007).
11. Talole, S., J. Kolhe, and S. Phadke, "Extended-state-observer-based control of flexible-joint system with experimental validation," *IEEE Trans. Ind. Electron.*, Vol. 57, pp. 1411–1419 (2010).
12. Ulrich, S. and J. Sasiadek, "Extended kalman filtering for flexible joint space robot control," *Proc. 2001 Am. Control Conf., San Francisco, California*, pp. 1021–1026 (2011).
13. Lee, D. J. and M. Tomizuka., "State/ parameter/ disturbance estimation with an accelerometer in precision motion control of a linear motor," *Proc. 2001 ASME Int. Mechatronics Eng. Congr. Expo., New York*, pp. 11–16 (2001).
14. Jeon, S. and M. Tomizuka, "Benefits of acceleration measurement in velocity estimation and motion control," *Control Eng. Practice*, Vol. 15, No. 3, pp. 325–332 (2007).
15. Chen, W. and M. Tomizuka, "Estimation of load side position in indirect drive robots by sensor fusion and kalman filtering," *Proc. 2010 Amer. Control Conf., Baltimore, Maryland*, pp. 6852–6857 (2010).
16. Franklin, G., J. Powell, and M. Workman, *Digital Control of Dynamic Systems*. Addison-Wesley, Boston, MA, second ed., (1990).
17. Ghorbel, F. H., P. S. Gandhi, and F. Alpeter, "On the kinematic error in harmonic drive gears," *J. Mech. Des.*, Vol. 123, No. 1, pp. 90–97 (2001).
18. Mehra, R., "Approaches to adaptive filtering," *IEEE Trans. Autom. Control*, Vol. 17, No. 5, pp. 693–698 (1972).

19. Maybeck, P. S., *Stochastic Models, Estimation, and Control*, Vol. 2. Academic Press, Inc., New York (1982).
20. Lambrechts, P., M. Boerlage, and M. Steinbuch, "Trajectory planning and feedforward design for electromechanical motion systems," *Control Eng. Practice*, Vol. 13, No. 2, pp. 145–157 (2005).
21. Chen, W. and M. Tomizuka, "Load side state estimation in robot with joint elasticity," *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics, Kaohsiung, Taiwan*, pp. 598–603 (2012).
22. Shumway, R. H. and D. S. Stoffer, "An approach to time series smoothing and forecasting using the em algorithm," *J. Time Ser. Anal.*, Vol. 3, pp. 253–264 (1982).
23. Ghahramani, Z. and G. E. Hinton, "Parameter estimation for linear dynamical systems," Tech. Rep. CRG-TR-96-2, University of Toronto (1996).

Wenjie Chen received B.Eng. degree from Zhejiang University, Zhejiang, China, in 2007, and the M.S. and Ph.D. degrees from University of California, Berkeley, CA, USA, in 2009 and 2012, respectively, all in Mechanical Engineering.

Dr. Chen is currently Postdoctoral Scholar in Department of Mechanical Engineering at the University of California, Berkeley. His research interests include design and implementation of advanced control and sensing algorithms with applications to robotic/mechatronic systems, such as industrial robots, wearable assistive robotics, and robots for advanced manufacturing.

Masayoshi Tomizuka received B.S. and M.S. degrees in Mechanical Engineering from Keio University, Tokyo, Japan, and Ph.D. degree in Mechanical Engineering from Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1974.

Dr. Tomizuka joined Department of Mechanical Engineering, University of California, Berkeley, in 1974, where he is currently the Cheryl and John Neerhout, Jr., Distinguished Professor. He teaches courses on dynamic systems and controls and conducts research on optimal and adaptive control, digital control, motion control, and their applications to robotics, manufacturing, information storage devices, and vehicles.

Dr. Tomizuka served as Program Director of Dynamic Systems and Control Program of the Civil and Mechanical Systems Division of National Science Foundation (2002–2004). He was Technical Editor of ASME Journal of Dynamic Systems, Measurement and Control (1988–1993), and Editor-in-Chief of IEEE/ASME Transactions on Mechatronics (1997–1999). He was the recipient of Rudolf Kalman Best Paper Award (ASME, 1995, 2010), DSCD Outstanding Investigator Award (ASME, 1996), Charles Russ Richards Memorial Award (ASME, 1997), Rufus Oldenburger Medal (ASME, 2002), and John R. Ragazzini Award (American Automatic Control Council, 2006). He is Fellow of American Society of Mechanical Engineers (ASME), the Institute of Electrical and Electronics Engineers (IEEE), the International Federation of Automatic Control (IFAC), and the Society of Manufacturing Engineers (SME).