Improving the accuracy of analog encoders via Kalman filtering

Article *in* Control Engineering Practice · April 2006

DOI: 10.1016/j.conengprac.2005.01.008

CITATIONS
11

READS
763

3 authors, including:

Yaakov Oshman
Technion - Israel Institute of Technology
155 PUBLICATIONS **2,298** CITATIONS

SEE PROFILE

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Control Engineering Practice 14 (2006) 337–350
CONTROL ENGINEERING PRACTICE
www.elsevier.com/locate/conengprac

Improving the accuracy of analog encoders via Kalman filtering ☆

Yaron Zimmerman[a,1], Yaakov Oshman[b,*], Amit Brandes[c,2]

[a]Technion—Israel Institute of Technology, Department of Mechanical Engineering, Technion City, Haifa 32000, Israel
[b]Technion—Israel Institute of Technology, Department of Aerospace Engineering, Technion City, Haifa 32000, Israel
[c]University of Rome 'La Sapienza', Computer and Systems Science Department, Via Eudossiana 18, 00184, Rome, Italy

Received 16 May 2004; accepted 24 January 2005
Available online 26 February 2005

Abstract

A new decoding method is presented for analog encoders enabling major improvements in both accuracy and resolution. A simulation study and experiments with real, industrial-grade, equipment demonstrate the performance improvement of the proposed method, revealing that the new method can generate position estimates with accuracy about three times better than that of standard methods. Moreover, in some special cases, the resulting position accuracy can reach sub-nanometer levels, thus enabling further size reduction in the semiconductor industry. The proposed algorithm also yields velocity estimates better by about two orders of magnitude than those obtained with standard methods.

Keywords: Encoders; Kalman filters; Estimation algorithms; Nonlinear filters; Accuracy; Measurement noise

1. Introduction

Linear analog encoder is a very common sensor in the semiconductor industry. Its main use is as a feedback sensor in X–Y or X–Y–Z tables that hold the wafer against a vision instrument like a high-resolution camera or an electronic microscope. In such applications the resolution requirements are extremely stringent, reaching few nanometers or even less than one nanometer. Existing decoding algorithms are limited because they are based on solving the geometrical problem only and do not attempt to improve the accuracy by applying some filtering strategy.

The analog encoder generates sine and cosine signals that are related to its linear translation. Like in any other real-world device, the encoder's measured signals are corrupted by common error sources, such as electronic noise or quantization error, which, in turn, reduce the accuracy of the encoder. The noise issue becomes even more important when bearing in mind that the encoder, a position sensor by its nature, is also commonly used to estimate velocities by a straightforward numerical differentiation of its output. To the best of the authors' knowledge, the idea of implementing state filtering techniques to confront the effects of the measurement error sources has not appeared in the open literature up to this date. Furthermore, only a few published works relate to the general problem of encoder measurement noise alleviation. Yang, Rees, and Chuter (2002) address only deterministic error sources, arising from mechanical installation errors, in an analog encoder. They develop a mathematical model of the resulting error and use a Kalman filter, implemented off-line, to estimate the model parameters. This model, in turn, is used to recalculate the measured position. The work of Venema (1994) is of a similar

☆Based on a paper presented at the 5th IFAC International Symposium on Intelligent Components and Instruments for Control Applications (SICICA-2003), Aveiro, Portugal, July 9–11, 2003.

*Corresponding author. Tel.:/fax: +97248293803.
E-mail addresses: yaron.zimmerman@kla-tencor.com (Y. Zimmerman), yaakov.oshman@technion.ac.il (Y. Oshman), amit_brandes@amat.com (A. Brandes).
[1]Currently with KLA-Tencor Corporation (Israel), Hatikshoret St., P.O.Box 143, Migdal Haemek 23100, Israel.
[2]Currently with Applied Materials, PDC business group, 9 Oppenheimer St., Rehovot 76705, Israel.

0967-0661/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.conengprac.2005.01.008

nature, using the filtering algorithm off-line for the sole purpose of calibrating the analog encoder. Other researchers (Hagiwara, Suzuki, & Murase, 1992; Mayer, 1994) deal only with the issue of calculating the encoder angle from the sine and cosine signal by applying inverse trigonometric functions.

This paper deals with the accuracy issue of analog encoders by applying on-line filtering to the measured signals. Using the extended Kalman filtering (EKF) framework, a novel decoding algorithm is developed and verified. The new algorithm enables improved position estimation and a dramatic improvement in velocity estimation, relative to commercially available existing algorithms.

The remainder of this paper is organized as follows. In the next section the operational principles of an analog encoder are described and the measurement model is constructed. The system model is then developed and both models are put together within the framework of an EKF. The following section presents a simulation study through which the filter's performance is verified, and a discussion of the results. Some of the more interesting and promising aspects of the proposed decoding method are pointed out. Next, the proposed algorithm is experimented with real industrial hardware and confronted with a conventional algorithm. Concluding remarks are offered in the final section.

2. Encoder model

2.1. Principle of operation

The encoder's principle of operation is schematically depicted in Fig. 1. A scale, having gold-made triangular facets, reflects the light of an infra-red LED light source, through an index grating, unto a photo-detector. Due to the periodic pattern of both the scale and the index grating, sinusoidal interference fringes are produced on the detection plane of the photo-detector. Whenever the moving parts are in motion, i.e., the light source, the index grating and the photo detector move relative to the static scale, the fringes move along the detection plane. Using a special assembly, the photo detectors provide signals related to the motion of the fringes. An electronic circuit amplifies and combines the photo detectors' signals to generate two sinusoidal waveforms of equal amplitude and period (equaling the scale period), separated by a 90 degrees phase shift. The interested reader is referred to (Webster, 1999; Renishaw, 1998) for further details.

2.2. Measurement model

The most accurate scales now commercially available have resolutions of a few micrometers. To obtain sub-micron resolution, as required by production systems, the analog signals must be sampled and decoded. The measurement vector $\boldsymbol{z} \in \mathbb{R}^2$ can be described as follows:

$$\boldsymbol{z} \triangleq \begin{bmatrix} z(1) \\ z(2) \end{bmatrix} = \begin{bmatrix} V\sin(2\pi f_s x(t)) \\ V\cos(2\pi f_s x(t)) \end{bmatrix}, \tag{1}$$

where f_s is the scale spatial frequency, V is the signal amplitude and $x(t)$ is the position.

Most manufacturers use the following equation to estimate the position from the measured signals (Venema, 1994):

$$x = \frac{a\tan2(z(1), z(2))}{2\pi f_s}, \tag{2}$$

where $a\tan2(\cdot,\cdot)$ is the 4 quadrant inverse tangent function. Some manufacturers increase the estimation accuracy by normalizing the signals and applying an arcsine or an arccosine function, depending on the signal absolute value, as follows:

$$x = \begin{cases} \dfrac{\arcsin(z(1)/\|\boldsymbol{z}\|)}{2\pi f_s}, & |z(1)| \leqslant |z(2)|, \\ \dfrac{\arccos(z(2)/\|\boldsymbol{z}\|)}{2\pi f_s}, & |z(2)| < |z(1)|. \end{cases} \tag{3}$$

This procedure increases the signal-to-noise ratio (SNR) by taking the information from the more sensitive regions in the signal harmonic.

All of the above algorithms do not apply any filtering procedure to the signal, i.e., they do not use any prior information regarding the signal to weigh it against the current measured information. In the presence of measurement noise, the algorithm's output, i.e., the estimated position, becomes noisy as well. Furthermore, the estimated velocity signal, calculated by numerically differentiating the (noisy) position signal, is subject to much higher noise levels.

To account for the sampling operation that transforms the continuous signal $x(t)$ to its sampled version $x(k+1)$, and for the contamination of the measurement by the measurement noise, Eq. (1) is rewritten as

$$\begin{aligned} \boldsymbol{z}(k+1) &= \boldsymbol{h}(x(k+1)) + \boldsymbol{v}(k+1) \\ &= \begin{bmatrix} V\sin(2\pi f_s x(k+1)) \\ V\cos(2\pi f_s x(k+1)) \end{bmatrix} + \boldsymbol{v}(k+1), \end{aligned} \tag{4}$$

where $\boldsymbol{v}(k+1)$ is the measurement noise. There are several sources for measurement noise (Venema, 1994), but the most dominant one is introduced by an electronic noise within the sampling operation (ground noise of the analog-to-digital (A/D) converter). The measurement noise is assumed to be a white, zero-mean Gaussian stationary sequence:

$$\boldsymbol{v} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}). \tag{5}$$

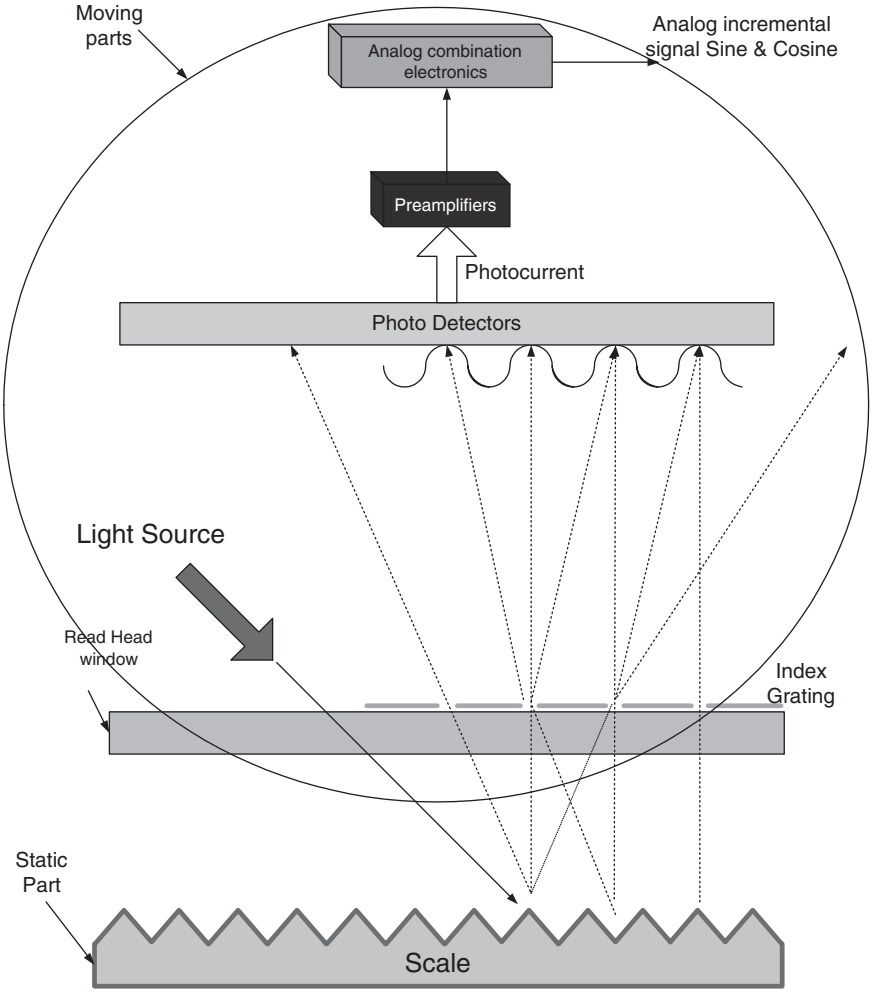

Fig. 1. Analog encoder principle of operation.

Eq. (4), relating the measurement vector $\boldsymbol{z}$ to the desired position signal x, is nonlinear. The filtering algorithm used is, therefore, the well-known EKF. The state equation, needed also in the EKF framework, is discussed next.

2.3. State space model

Two basic concepts for the state space model are possible:

(1) a full dynamic model, which takes into account the mechanical structure, material property, actuators and sensors location and actuated forces,
(2) a kinematic model, which take into account just the kinematic relation between the state variables.

Although it is conceivable that the full dynamic model would yield better estimation performance than the kinematic model because of its better description of the true system, the latter model was chosen in the present work because of the complexity of the dynamic model and its inherent sensitivity to the quality of its underlying mathematical model.

Following Bar-Shalom and Fortmann (1988) a kinematic model of a linear system driven by a white noise was adopted. The state variables are the position, the velocity and the acceleration of the system. The state vector is, thus

$$\boldsymbol{x} = [x \;\; \dot{x} \;\; \ddot{x}]^{\mathrm{T}} \tag{6}$$

and the state equation is

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\,\boldsymbol{x} + \boldsymbol{w} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\alpha \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix}, \tag{7}$$

where w is assumed to be a white, zero-mean Gaussian process with intensity σ_m^2, and α is the inverse decorrelation time (pole) of the acceleration. Eqs. (4) and (7) comprise the full (nonlinear) continuous model of the analog encoder, to be used in the sequel as a basis for the EKF formulation.

3. EKF implementation

Similar to (Bar-Shalom & Fortmann, 1988) a discrete-time model of the system is derived as follows:

$$\boldsymbol{x}(k+1) = \boldsymbol{F}\,\boldsymbol{x}(k) + \boldsymbol{w}_d(k), \tag{8}$$

where the state transition matrix is

$$\boldsymbol{F} = \mathrm{e}^{\boldsymbol{A}\,T} = \begin{bmatrix} 1 & T & \dfrac{\alpha T - 1 + \mathrm{e}^{-\alpha T}}{\alpha^2} \\ 0 & 1 & \dfrac{1 - \mathrm{e}^{-\alpha T}}{\alpha} \\ 0 & 0 & \mathrm{e}^{-\alpha T} \end{bmatrix} \tag{9}$$

and T is the sampling time. The symmetric covariance matrix $\boldsymbol{Q}_d$ of the discrete-time process noise $\boldsymbol{w}_d$ has the following components:

$$q_{11} = \frac{\sigma_m^2}{\alpha^4}\left[1 - \mathrm{e}^{-2\alpha T} + 2\alpha T + \frac{2\alpha^3 T^3}{3} - 2\alpha^2 T^2 - 4\alpha T \mathrm{e}^{-\alpha T}\right], \tag{10a}$$

$$q_{12} = \frac{\sigma_m^2}{\alpha^3}[\mathrm{e}^{-2\alpha T} + 1 - 2\mathrm{e}^{-\alpha T} + 2\alpha T \mathrm{e}^{-\alpha T} - 2\alpha T + \alpha^2 T^2], \tag{10b}$$

$$q_{13} = \frac{\sigma_m^2}{\alpha^2}[1 - \mathrm{e}^{-2\alpha T} - 2\alpha T \mathrm{e}^{-\alpha T}], \tag{10c}$$

$$q_{22} = \frac{\sigma_m^2}{\alpha^2}[4\mathrm{e}^{-\alpha T} - 3 - \mathrm{e}^{-2\alpha T} + 2\alpha T], \tag{10d}$$

$$q_{23} = \frac{\sigma_m^2}{\alpha}[\mathrm{e}^{-2\alpha T} + 1 - 2\mathrm{e}^{-\alpha T}], \tag{10e}$$

$$q_{33} = \sigma_m^2[1 - \mathrm{e}^{-2\alpha T}]. \tag{10f}$$

The EKF mechanization equations are standard and can be found in many textbooks, e.g., (Mendel, 1987). Let $\hat{\boldsymbol{x}}(k+1\,|\,k)$ denote the predicted estimate at time t_{k+1} based on k measurements, and let $\boldsymbol{H}_x$ denote the Jacobian matrix of the measurement function, calculated at the predicted state:

$$\boldsymbol{H}_x(k+1) \triangleq \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}(k+1\,|\,k)} = \begin{bmatrix} 2V\pi f_s \cos(\hat{x}(k+1\,|\,k)) & 0 & 0 \\ -2V\pi f_s \sin(\hat{x}(k+1\,|\,k)) & 0 & 0 \end{bmatrix}. \tag{11}$$

The EKF is then formulated as follows:

- Time update:

$$\hat{\boldsymbol{x}}(k+1\,|\,k) = \boldsymbol{F}\,\hat{\boldsymbol{x}}(k\,|\,k), \tag{12}$$

$$\boldsymbol{P}(k+1\,|\,k) = \boldsymbol{F}\,\boldsymbol{P}(k\,|\,k)\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{Q}_d. \tag{13}$$

- Measurement update:

$$\hat{\boldsymbol{x}}(k+1\,|\,k+1) = \hat{\boldsymbol{x}}(k+1\,|\,k) + \boldsymbol{K}(k+1)[\boldsymbol{z}(k+1) - \boldsymbol{h}(\hat{\boldsymbol{x}}(k+1\,|\,k))], \tag{14}$$

$$\boldsymbol{K}(k+1) = \boldsymbol{P}(k+1\,|\,k)\boldsymbol{H}_x(k+1)^{\mathrm{T}}[\boldsymbol{H}_x(k+1)\boldsymbol{P}(k+1\,|\,k)\boldsymbol{H}_x(k+1)^{\mathrm{T}} + \boldsymbol{R}]^{-1}, \tag{15}$$

$$\boldsymbol{P}(k+1\,|\,k+1) = [\boldsymbol{I} - \boldsymbol{K}(k+1)\boldsymbol{H}_x(k+1)]\boldsymbol{P}(k+1\,|\,k)[\boldsymbol{I} - \boldsymbol{K}(k+1)\boldsymbol{H}_x(k+1)]^{\mathrm{T}} + \boldsymbol{K}(k+1)\boldsymbol{R}\,\boldsymbol{K}(k+1)^{\mathrm{T}}. \tag{16}$$

4. Simulation study

A MATLAB simulation was constructed to evaluate the performance of the proposed algorithm and to compare it with that of standard decoding methods. The simulation, described schematically in Fig. 2, consists of a signal generator that generates the true position signal, an encoder block, a sampler (A/D) and a processing block that implements either the proposed algorithm or a standard decoding method, based on inverse-trigonometric functions.

The encoder block simulates the sine and cosine signals of an analog encoder according to a preset scale resolution. The signals are then sampled by the sampler block that enables also setting of the converter resolution and of the electronic noise level, simulated as an additive Gaussian white noise. The parameters (dynamic range, resolution) of the analog-to-digital converter are set in accordance with standard commercial components. These parameters are detailed in Table 1. The electronic noise level was specified by a leading manufacturer of X–Y table controllers and verified with real data.

The processing block comprises two different functions. The 'coarse position' function counts full cycles of the encoder scale and is identical for the standard and the proposed algorithms. The 'fine position' function is the heart of the decoding process and calculates the position within the scale cycle according to the sine and cosine signals of the encoder. This function implements either the standard inverse-trigonometric algorithm of Eq. (3), or the proposed EKF-based algorithm.

The tuning of an EKF is a delicate matter regarded by many engineers as an art more than a science. Two tuning parameters exist in this case: the process noise intensity σ_m^2 and the inverse decorrelation time α. Both parameters control the bandwidth of the filter. The

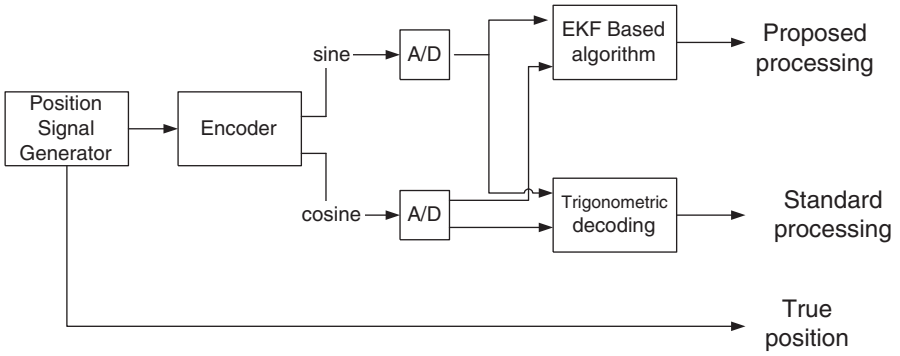

Fig. 2. Simulation block diagram.

Table 1
Simulation parameters

Parameter	Value
Sampler dynamic range (V)	±1.25
Sampler resolution (bits)	12
Sampling frequency (kHz)	20
Signal amplitude (V)	1
Encoder scale cycle (μm)	4
Electronic noise 1σ (mV)	8
Acceleration pole (α) (rad/s)	200π
Process noise intensity (σ_m^2) $((m/s^3)^2)$	10^{-5}

following considerations are taken into account when tuning the EKF:

- The filter's bandwidth has to be tuned to match the expected signal frequencies;
- High filter bandwidth results, typically, in noisier estimates and larger estimation error, whereas low bandwidth can result in bad tracking performance (higher estimation delay and, even, filter divergence);
- Precautions should be taken to ensure sufficient signal power to account for the expected accelerations in the system.

In general, lower values of the σ_m^2 parameter result in lower values of the $\boldsymbol{P}$ matrix. But, if set too low, then the estimation becomes delayed and the estimation error tends to become sinusoidal (for sinusoidal inputs). Lower α values enable setting lower σ_m^2 values, thus permitting a lower error standard deviation, but at the expense of lowering the bandwidth of the filter.

Having in mind a closed loop bandwidth of about 10–20 Hz, as is common in industrial X–Y tables, the filter was tuned with $\alpha = 200\pi$ to account for acceleration bandwidth of at least 100 Hz. The filter was tuned to perform well with signals having amplitudes on the order of a few microns, which represent normal requirements for such systems. The sampling frequency was set to 20 kHz. The performance was compared to that of the standard inverse trigonometric method by calculating the point-wise position and velocity errors and the standard deviations of the errors for both algorithms.

The algorithms are tested either by generated sinusoidal position inputs with various frequencies, or by position ramp (constant velocity) inputs. These signals are different than those assumed in the kinematic model that lies in the basis of the EKF. Moreover, the model assumes only additive measurement noise whereas the simulated measurements are affected by sampling, quantization and an additive random noise. The estimation errors, generated by the two decoding algorithms, are compared for the various input signals. In addition, the new method's estimation error standard deviation is compared to the theoretical value computed from steady state values of the filter covariance matrix, to verify filter consistency.

Fig. 3 presents the true position signal vs. the estimated one, on a magnified time scale. Fig. 4 shows the difference between those signals, i.e., the estimation error of the two decoding algorithms. These figures show the improvement in position estimation due to the use of the new decoding algorithm. When using the proposed algorithm the errors are clearly smaller and the high-frequency (noisy) components are significantly attenuated. Table 2 presents the standard deviation of the position error generated by the two decoding algorithms, for various driving signals. Remarkably, the new method's error standard deviation is kept close to 1.5 nm even when the amplitude or the frequency of the position signal is varied. Tracking of a constant velocity motion is also handled well by the filter, with the same level of accuracy.

A dramatic improvement is achieved in the velocity estimation, as could be expected. This is mainly due to the fact that standard methods estimate velocity by numerically differentiating the noisy position estimate, whereas the proposed algorithm estimates it as part of the state vector. The results shown in Table 3 and in Fig. 5 clearly demonstrate the huge improvement in velocity estimation (note the different scales of the vertical axes in Fig. 5).

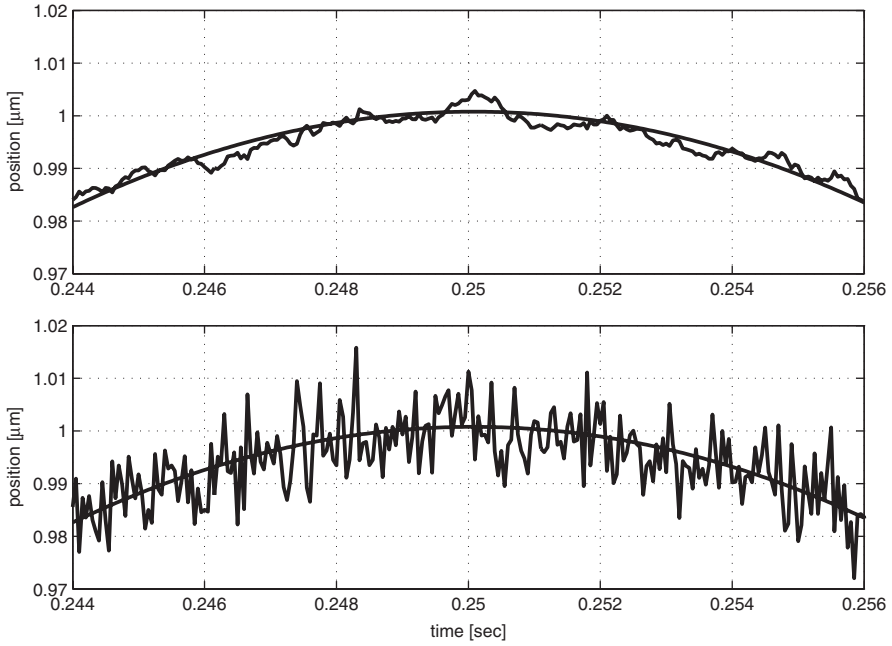

Fig. 3. True vs. estimated position, EKF (top) and inverse trigonometric (bottom) decoding methods.

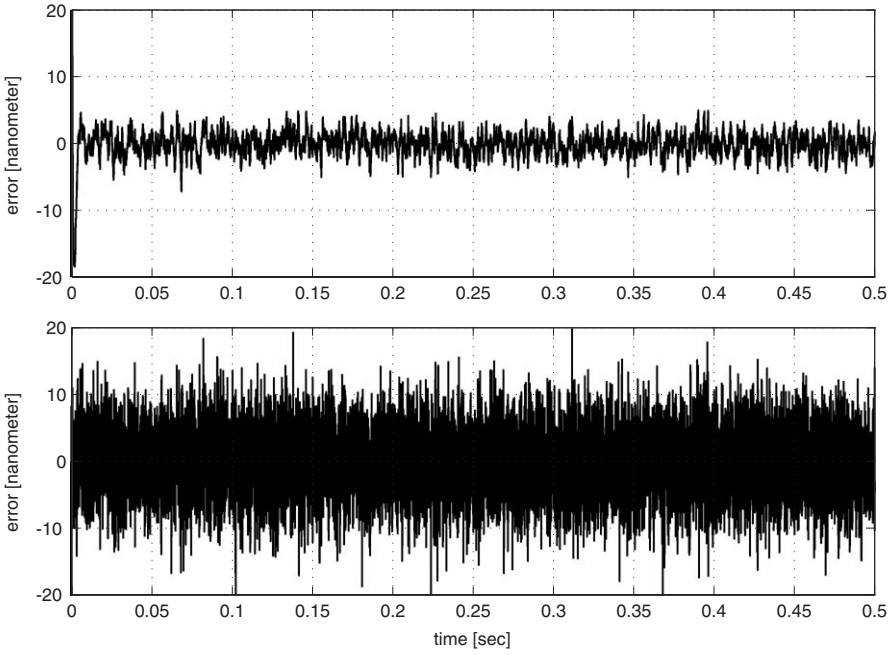

Fig. 4. Position estimation error, EKF (top) and inverse trigonometric (bottom) decoding methods.

Table 2
Position 1-σ error (nanometer) for sinusoidal signals (designated by frequency and amplitude) and constant velocity signal (designated by amplitude)

Decoding method	5 Hz 1 µm	2 Hz 1 µm	10 Hz 1 µm	5 Hz 2 µm	5 Hz 0.5 µm	Constant velocity 2 mm/s
Inverse trigonometric	5.17	5.15	5.15	5.16	5.15	5.13
EKF	1.57	1.54	1.67	1.54	1.56	1.53

Table 3
Velocity 1-σ error (μm/s) for sinusoidal signals (designated by frequency and amplitude) and a constant velocity signal (designated by amplitude)

Decoding method	5 Hz 1 μm	2 Hz 1 μm	10 Hz 1 μm	5 Hz 2 μm	5 Hz 0.5 μm	Constant velocity 2 mm/s
Inverse trigonometric	146.2	145.4	144.9	145.1	145.4	144.5
EKF	1.77	1.69	2.29	1.84	1.72	1.69

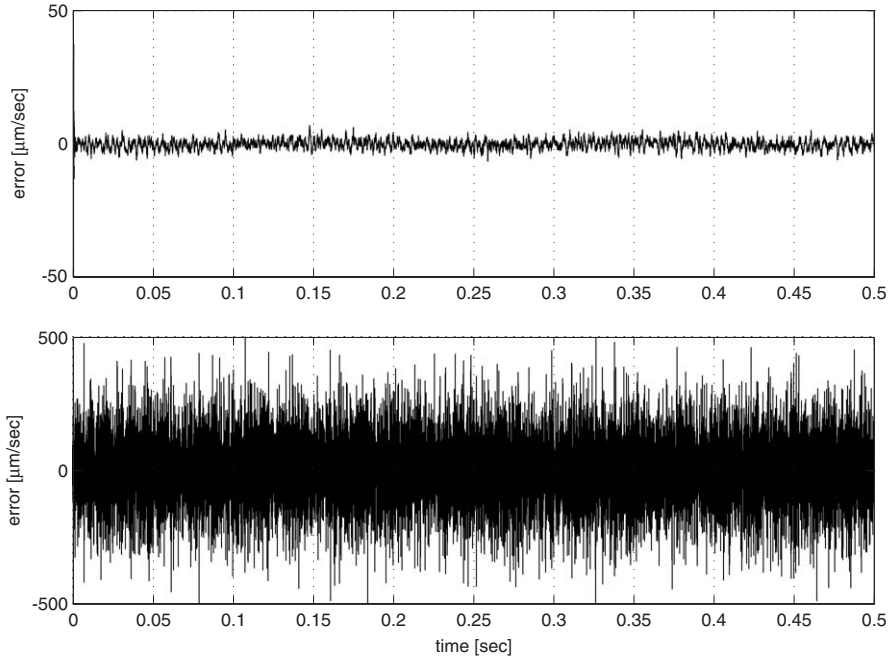

Fig. 5. Velocity estimation error, EKF (top) and inverse trigonometric (bottom) decoding methods (notice the different vertical scales).

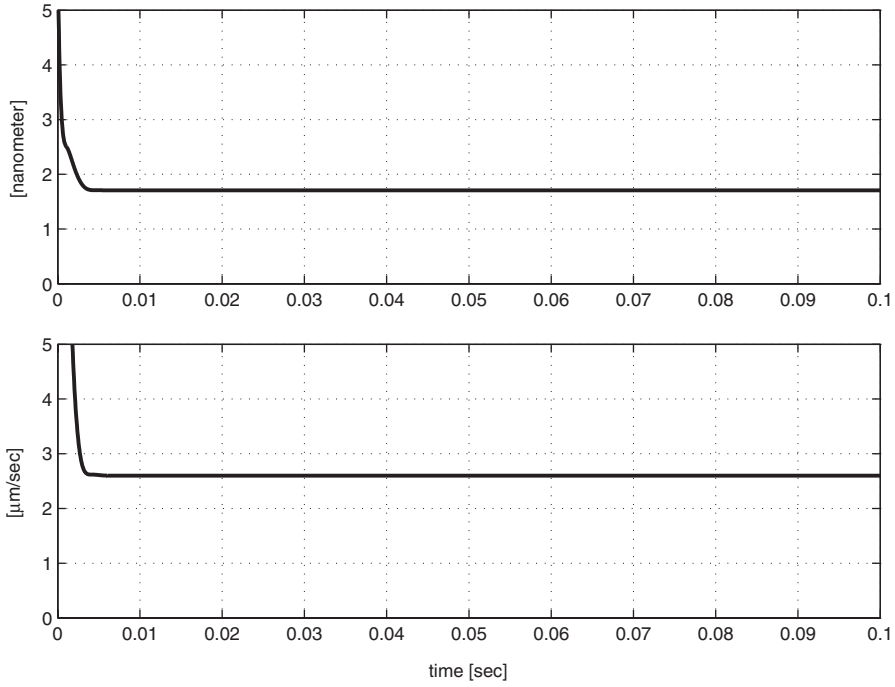

Fig. 6. Predicted error standard deviation calculated from the EKF covariance matrix. Position (top) and velocity (bottom).

Fig. 6 presents the predicted error standard deviation as calculated from the EKF error covariance matrix. Note the good correlation of the accuracy predicted by the filter, 1.7 nm in position and 2.6 μm/s in velocity, with the actual results shown in Tables 2 and 3. Note also the fast convergence of the filter and its quasi-static behavior.

In some special cases, wafer production and inspection systems are subject to very low amplitudes in position (quasi-static behavior). When simulating such cases results of sub-nanometer position accuracies were obtained by using the proposed algorithm. This is mainly due to the fact that the low values of acceleration enable tuning of the filter with relatively low values of σ_m^2, as shown in Table 4. Accuracy is also commonly improved by raising the sampling frequency. To-date commercial controllers are limited to sampling frequency of about 20 kHz, but future hardware development will surely enable higher rates. Table 4 demonstrates that sub-nanometer accuracy can be

Table 4
Simulation results—sub-nanometer accuracy, 2 Hz signals (position and velocity errors are 1-σ values)

Signal amplitude (μm)	σ_m^2 $((m/s^3)^2)$	Frequency (kHz)	Position error (nm)	Velocity error (μm/s)
0.1	10^{-9}	20	0.6	0.08
1	10^{-5}	100	0.81	1.22
1	10^{-5}	20	1.54	1.69

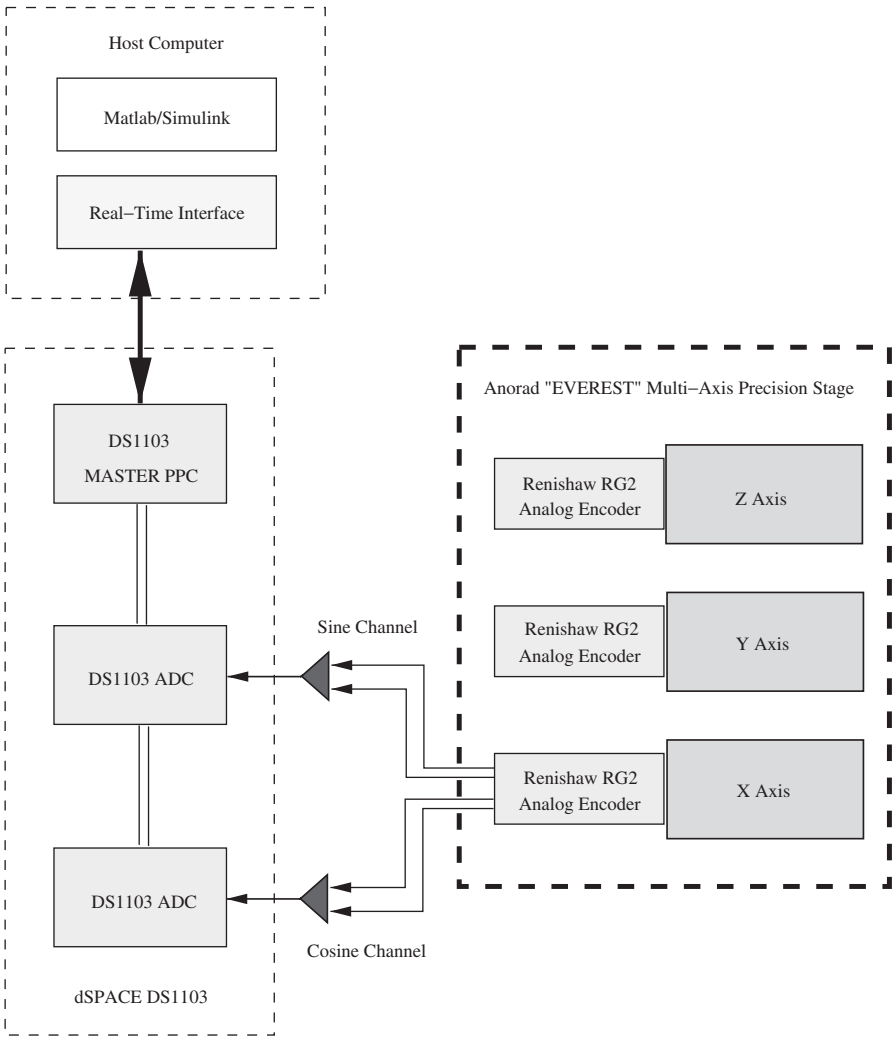

Fig. 7. Schematic diagram of experimental setup.

achieved by raising the sampling frequency to 100 kHz. The two cases are compared to the normal setting of the filter (third row of Table 4). Note that the accuracy of the standard inverse trigonometric decoding method is independent of the above setting and remains ca. 5 nm.

5. Experimental validation

In the previous section the performance improvements associated with the new algorithm were demonstrated via a simulation study. To further validate the simulation and demonstrate the capability of the method, an experiment using real data, acquired via an operative industrial system, has been carried out. The experiment was performed at the Mechatronics Laboratory of the Technion—Israel Institute of Technology's faculty of mechanical engineering. The testbed centers on an "EVEREST" industrial X–Y–Z precision stage assembled by Anorad Ltd., using RG-2 linear analog encoders made by Renishaw Ltd., as the position sensors. The stage is normally used for wafer inspection, which requires resolution of 40 nm. A dSPACE data acquisition system with real-time interface (dSPACE, 2001, 2003) was connected to the X-axis encoder, thus enabling precise data recording. Fig. 7 depicts schematically the experimental setup. The hardware parameters are detailed in Table 5. Fig. 8 shows a picture of the Anorad Ltd. "EVEREST" multi-axis precision stage.

Table 5
Experimental setup parameters

Parameter	Value
Sampler dynamic range (V)	±10
Sampler resolution (bits)	12
Sampling frequency (kHz)	20
Signal amplitude (V)	5.45
Encoder scale cycle (μm)	20

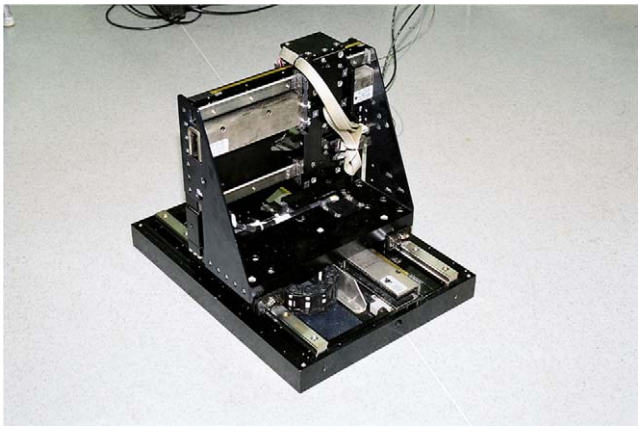

Fig. 8. Anorad-assembled "EVEREST" multi-axis precision stage used for the experimental investigation.

The experiments consisted of applying a series of motion profiles to the X-axis of the testbed. The resulting encoder signals were recorded and the proposed algorithm applied to the acquired data. The results were compared to the ones obtained when applying a conventional algorithm to the same data. The recorded data were first analyzed to determine noise parameters, important for a successful implementation of the proposed algorithm.

The experiments comprised the following two stages:

(1) static tests, in which the X-axis was held in some constant position,
(2) dynamic tests, where several dynamic motion profiles were executed.

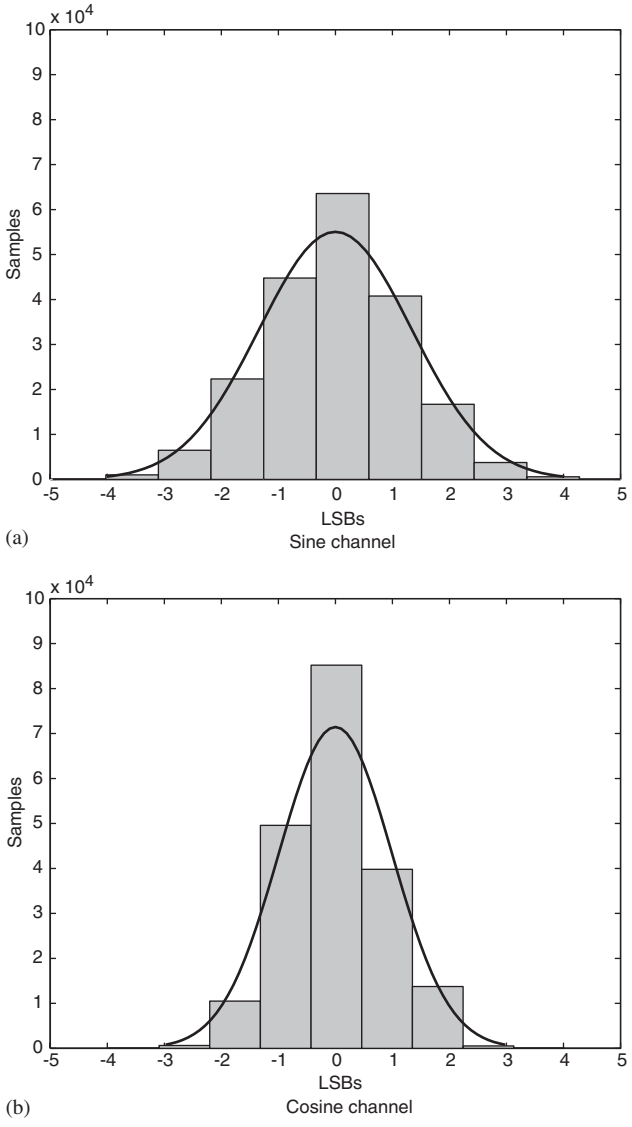

Fig. 9. Noise histograms and corresponding Gaussian PDFs in both encoder channels: (a) Sine channel; (b) cosine channel.

The former tests facilitate an assessment of the noise statistical properties due to the table's known constant position and zero velocity, both serving as reference signals for comparison. The latter tests were executed in order to verify and demonstrate the ability of the proposed algorithm to track dynamic motion profiles.

5.1. Noise statistics assessment

The statistical parameters of the measurement noise associated with each encoder channel (sine, cosine) were assessed from data acquired in the static tests, where all variations from the mean constant position are assumed

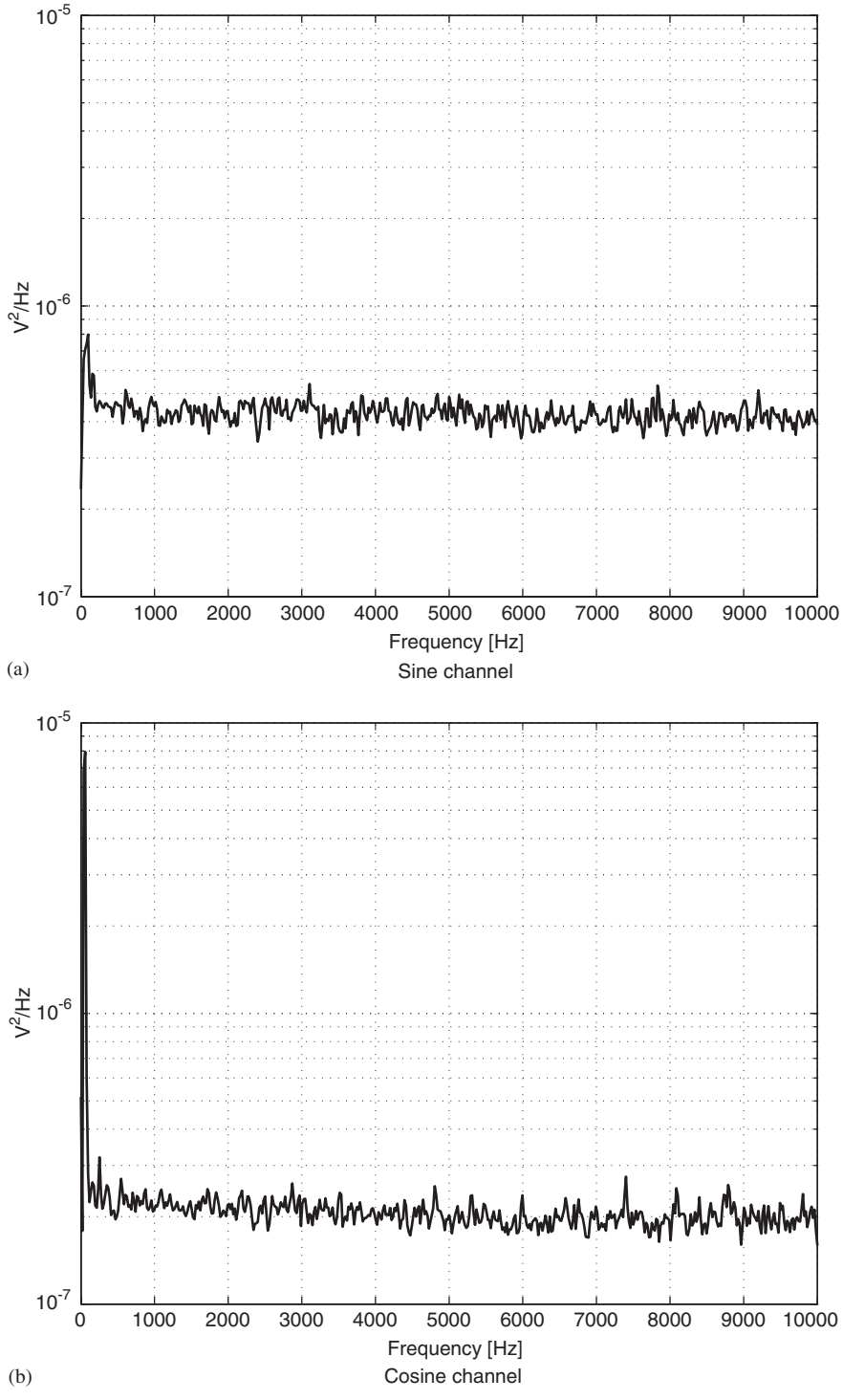

Fig. 10. Experimental PSDs of measurement noise in both encoder channels: (a) Sine channel; (b) cosine channel.

to be noise. Assuming ergodicity of the noise process, a qualitative assessment of the noise distribution in each encoder channel is obtained by plotting the noise histogram, giving a rough estimate of the first-order probability density function (PDF) of the process, and comparing it with the corresponding Gaussian PDF having the same mean and variance. The measured histograms and the corresponding normal PDFs thus obtained are compared in Fig. 9. The noise histograms are in good qualitative match with the corresponding Gaussian PDFs, thus justifying the hypotheses made in Section 2. Computed using the measured data, the noise standard deviations were found to be 1.5 LSBs (7.5 mV) for the sine channel and 1 LSB (5 mV) for the cosine channel of the encoder. These values were used in the noise covariance matrix of the EKF. Notice that these values are in good correspondence with the simulation parameters (see Table 1).

The power spectral densities (PSDs) of the sine and cosine signals were also computed, to qualitatively verify the whiteness of the signals. Fig. 10 depicts the respective PSDs. In general, both experimentally derived PSD functions demonstrate excellent consistency with the standard Kalman filter's white noise assumption, except for a 50 Hz peak, more evident in the cosine signal, which clearly corresponds to the 50 Hz alternating current signal in the power lines. Note also that the cosine channel's PSD is lower than that of the sine channel, which is in agreement with the respective standard deviations.

5.2. Static experiment

In the next stage of the experiment, the proposed algorithm, tuned to perform well in both static and dynamic conditions, was applied to the data recorded during the static tests. The results are compared to the reference position, taken as the mean of the recorded (noisy) readings and the reference velocity, taken to be zero in this stage. Note that the hardware parameters are different than the ones in the simulation section (see Table 1 vs. Table 5). In particular, the encoder scale in the simulation is five times smaller than that of the experimental setting, calling for less accurate results. Nevertheless, the accuracy improvement, in both position and velocity estimates, is clearly evident.

Fig. 11 shows the position estimation errors obtained by both the standard and the proposed decoding methods. Lacking an independent, sufficiently accurate measurement of the position that could serve as a reference for this test, the mean value of the measured position is taken as representing the true position. As can be seen from Fig. 11, the position estimate generated by the EKF-based method is much smoother than the standard method's corresponding result. The position estimation error standard deviation is reduced by the new method from a value of 6.6 nm, corresponding to the standard algorithm, to 2.4 nm. The improvement is almost threefold, as predicted by the simulation in Section 4.

Fig. 12 presents the corresponding comparison for the velocity estimates. The value of the true velocity is taken

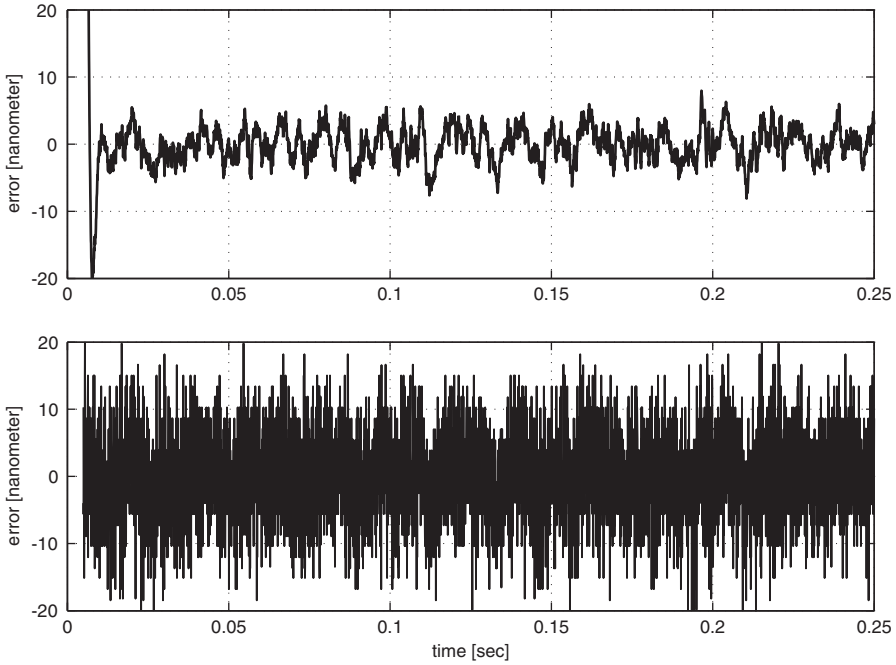

Fig. 11. Position estimation error in the static experiment: EKF (top) vs. inverse trigonometric (bottom) decoding methods.

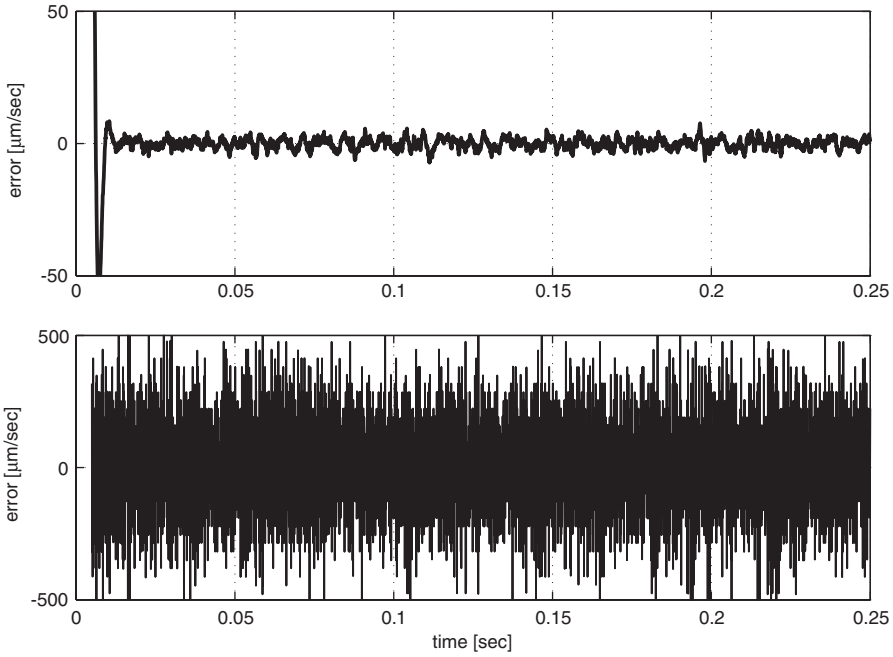

Fig. 12. Velocity estimation error in the static experiment: EKF (top) vs. inverse trigonometric (bottom) decoding methods (notice the different vertical scales).

Table 6
Position and velocity 1-σ estimation errors in static experiment

Decoding method	Position error (nm)	Velocity error (μm/s)
Inverse trigonometric	6.6	182.0
EKF	2.4	2.0

as zero in the static test, which serves as the reference to which both estimates are compared. As can be seen from Fig. 12, the velocity estimation error is reduced by the new method by more than an order of magnitude (notice the different scales of the vertical axes). The estimation error standard deviation corresponding to the new method is 2.0 μm/s, vs. a value of 182.0 μm/s corresponding to the inverse trigonometric decoding method. Again, these values are in good agreement with the results of the simulations presented in Section 4. Table 6 summarizes the position and velocity $1-\sigma$ errors of both algorithms in the static experiment.

5.3. Dynamic experiment results

In the last stage of the experiment, the new estimation algorithm was applied to data acquired when the table was in motion, in order to verify the algorithm's ability to track dynamic motion profiles. The same tuning parameters used during the static tests were employed during this stage. A typical motion profile is presented in Fig. 13. Fig. 14 presents estimates of this motion profile as computed by both the proposed EKF-based decoding

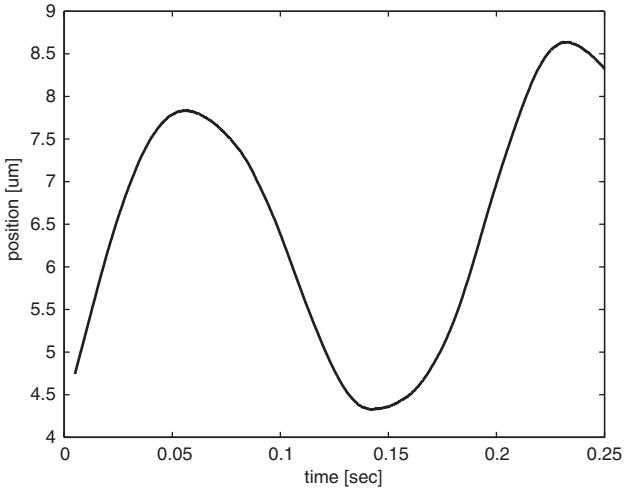

Fig. 13. A typical motion profile executed during the dynamic experiment.

method and the conventional one. As can be clearly observed, the estimates obtained using the proposed algorithm are much smoother than those obtained using the standard method. Although the accuracy of the estimates cannot be quantitatively determined in this test due to the absence of an independent, accurate reference signal, it can be safely stated that the smooth position and velocity estimates computed by the new method are in much better agreement with the table's dynamics, which is low-pass in nature.

Finally, Fig. 15 presents the velocity estimates during a typical dynamic motion test. As evident from Fig. 15, the standard algorithm generates an extremely coarse

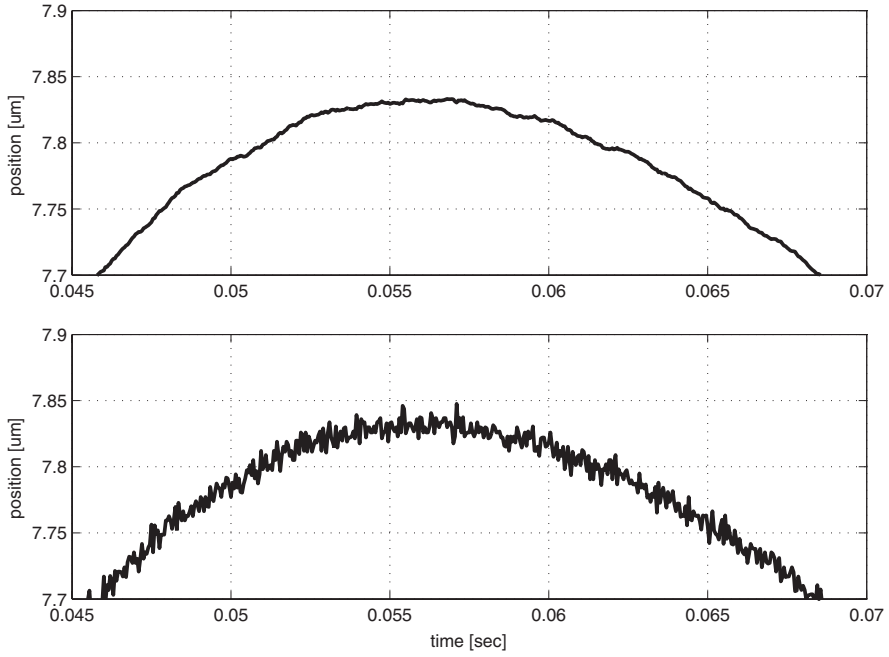

Fig. 14. Position estimates as computed by the new method (top figure) and the standard algorithm (bottom figure) during the dynamic experiment.

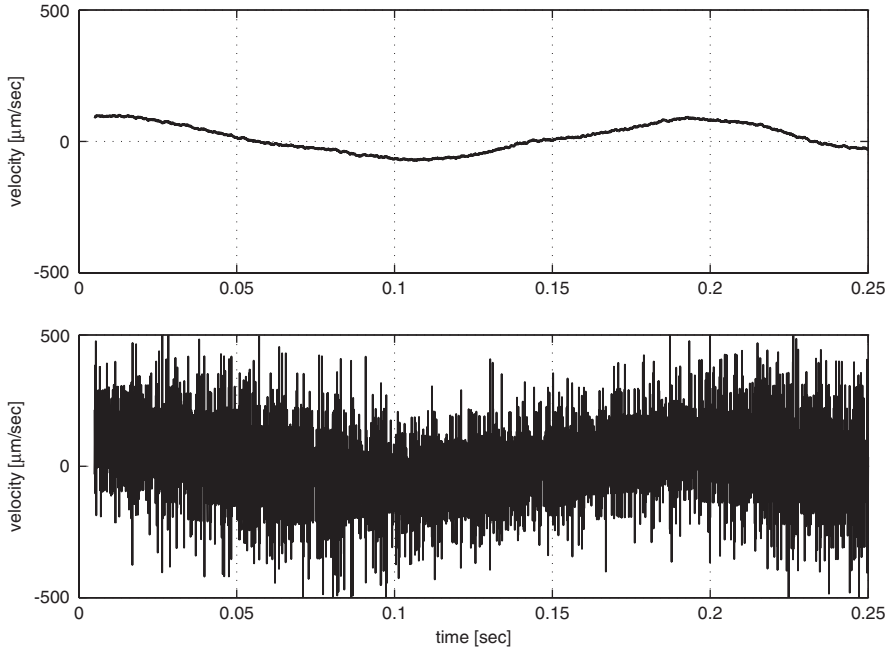

Fig. 15. Velocity estimates as computed by the new method (top figure) and the standard algorithm (bottom figure) during dynamic motion of the table.

velocity estimate, which is virtually useless for the purpose of driving a feedback control loop. In clear contradistinction, the newly proposed method generates a much smoother estimate that can be considered a good source for a control feedback signal, without presenting too much risk of loop stability. It should be noted that although low-pass pre-filtering of the velocity estimate is possible, the amount of filtering needed to smoothen the conventional algorithm's velocity estimate is so large that it is bound to introduce significant delays.

6. Conclusions

A new decoding method is presented for analog encoders. Based on the EKF technique, the new method

enables major improvements in accuracy and resolution. The algorithm can be easily implemented in multi-axis table controllers and other applications using analog encoders (linear or otherwise). The position accuracy is normally three times better than that of a standard inverse trigonometric decoding method and, in some special cases, can reach sub-nanometer levels, thus enabling further size reduction in the semiconductor industry while using off-the-shelve, commercially available sensors and hardware. Another important benefit of the proposed algorithm is its inherent velocity estimation capability, with velocity estimates better by two orders of magnitude than those obtained with standard methods. Guidelines regarding the filter tuning are also provided. Experiments using industrial hardware show good correspondence with simulated results and further point out the possibilities enabled using the proposed algorithm.

Acknowledgements

The authors wish to express their gratitude to the Anorad Ltd. and Mega-F Motion Systems Ltd. companies, that facilitated the experiment by generously providing some of the required experimental setup components. In particular, the help of the following individuals is gratefully acknowledged: Alex Novikov of Anorad Ltd., and Dr. Max Sakhartov and Avraham Levkovich of Mega-F Motion Systems Ltd. The authors also thank Dr. Izhak Bucher of the Technion—Israel Institute of Technology for providing the facilities and instrumentation of the Mechatronics Laboratory.

References

Bar-Shalom, Y., & Fortmann, T. E. (1988). *Tracking and data association*. San Diego, CA: Academic Press, Inc.

dSPACE (2001). *Real time interface implementation guide*. Paderborn: dSPACE.

dSPACE (2003). *dSPACE DS1103 PPC controller board installation and configuration guide*. Paderborn: dSPACE.

Hagiwara, N., Suzuki, Y., & Murase, H. (1992). A method of improving the resolution and accuracy of rotary encoders using a code compensation technique. *IEEE Transactions on Instrumentation and Measurements*, *41*(1), 98–101.

Mayer, J. R. R. (1994). High-resolution of rotary encodes analog quadrature signals. *IEEE Transactions on Instrumentation and Measurements*, *43*(3), 494–498.

Mendel, J. M. (1987). *Lessons in digital estimation theory*. Englewood Cliffs, NJ, USA: Prentice-Hall Inc.

Renishaw. (1998). *RG2 linear encoder system*. Renishaw Ltd., commercial catalog.

Venema, S. C. (1994). *A Kalman filter calibration method for analog quadrature position encoders*. Master's thesis, University of Washington, Seattle, Washington, USA.

Webster, J. (Ed.). (1999). *The measurement instrumentation and sensor handbook*. Boca Raton, FL, USA: CRC press.

Yang, Y., Rees, N., & Chuter, T. (2002). Reduction of encoder measurement errors in UKIRT telescope control system using a Kalman filter. *IEEE Transactions on Control Systems Technology*, *10*(1), 149–157.