Nonlinear Bounded State Estimation for Sensorless Control of an Electromagnetic Device

Eduardo Moya-Lasheras, Carlos Sagues, Edgar Ramirez-Laboreo and Sergio Llorente

Abstract—This paper presents a novel nonlinear state observer with discrete-time measurements for estimating the plunger position of linear travel solenoid valves. The observer is an unscented Kalman filter (UKF) for nonlinear systems that iteratively calculates an estimated mean and covariance of the state. It is based on a basic lumped parameter model, which contributes to the computational efficiency of the observer and facilitates its implementation. The magnetic reluctance is modeled taking into account the magnetic saturation and is partly defined by data obtained from finite element analysis (FEA). Boundary constraints are added to the estimated position to prevent it from surpassing its physical limits. Different tests performed with simulated and experimental data show that the estimations are accurate and robust to noise and model inaccuracies. Besides, although the observer has been developed for a specific device, the method can be easily extended to other electromechanical systems in which the position needs to be estimated.

I. INTRODUCTION

Electromagnetic actuators are very popular for a wide range of fields. While contactors and relays are typically utilized for power switching operations [1], [2], solenoid valves are used for flow control in many applications, e.g., in automotive internal combustion engines [3], [4] and electronic-stability control [5], [6]. Low-cost fast switching solenoid valves are typically used as safety devices, but an effective control strategy would permit using them in other applications, replacing more expensive valves. One of the major drawbacks of these devices is the high impact velocity reached in their opening and closing, which causes significant noise and wear. Therefore, the control strategy should achieve soft landing [3]. In order to implement a feedback control, the plunger position is needed, but note that there is no viable and affordable method to directly measure it.

A solution to this problem could be to design a nonlinear state observer from measurements of the coil voltage and current. A possibility is a sliding mode observer, which has been proposed e.g., by [7]. However, this strategy does not take into account the errors in the model and the measurements. Besides, the chattering that appears in systems with fast dynamics is difficult to eliminate and makes it undesirable for soft landing control. Another possibility is a Kalman filter. The original Kalman filter [8] predicts the internal state of a process given noisy measurements and an accurate model, but it works only for linear systems. The extended Kalman filter (EKF) can be applied to nonlinear systems by linearizing around the predicted state, and has been proposed for similar devices, such as electromagnetic valve actuators in internal combustion engines [4]. However, this approach can introduce large errors for highly nonlinear systems with non-Gaussian errors and model uncertainties.

In this paper we propose an estimator of the position based on the UKF [9], based on a nonlinear discrete-time model of a particular valve. The state of the system is iteratively estimated, including its position, along with its uncertainty expressed as an error covariance matrix. The robustness of the UKF to measurement errors and model uncertainties is analyzed by using simulated data, obtained from a hybrid automaton, and measured data.

The main contributions of this paper are: an application of the UKF algorithm for online position estimation of a low-cost fast switching electromechanical device; a system representation with a hybrid automaton which accounts for the velocity discontinuities during contact, adding white noise to analyze the UKF estimation; and a reluctance model which is divided into a parametric part that accounts for the magnetic saturation of the core, and a lookup table that depends on the gap length.

II. SYSTEM MODELING

The model represents a particular linear travel solenoid valve, shown in Fig. 1. It has a cylindrically symmetrical steel core with a fixed and a movable part. The steel housing permits the magnetic flux to return. The coil is wrapped between the core and the housing. A helical spring keeps the two parts of the core separated when the coil is de-energized.

For online state estimation, the observer needs to be based on a computationally fast model. FEA models are accurate, but time-consuming. On the other hand, lumped parameter models have low computing requirements. However, their simplifications may lead to inaccuracies, especially in the magnetic reluctance and force. For these reasons, the system is represented with a lumped parameter model based on electrical and magnetic circuits that also relies on data obtained from FEA simulations.
The system can be divided into its electromagnetic and mechanical subsystems.

A. Electromagnetic subsystem

The electrical equation of the coil is

\[ v = Ri + N \frac{d\phi}{dt}, \]

where \( v \) is the voltage between the coil terminals, \( i \) is the current through the coil, \( R \) is the internal resistance that accounts for Joule losses, \( N \) is the number of turns of the coil and \( \phi \) is the magnetic flux generated by the coil current.

Additionally, the coil current and the magnetic flux can be interrelated by Hopkinson's law,

\[ Ni = R\phi, \]

where \( R \) is the magnetic reluctance. It can be expressed as an integral depending on the geometry and on the material permeability \( \mu \) of the valve,

\[ R = \int \frac{dl}{\mu(l) A(l)}, \]

where the variable \( l \) is the position on the path of the magnetic flux and \( A \) is the cross-sectional area. The reluctance can be divided into two terms corresponding to the core and the gap,

\[ R = \int \frac{dl_c}{\mu_c(l_c) A(l_c)} + \int \frac{dl_g}{\mu_0 A(l_g)} = R_c + R_g, \]

being \( R_c \) the reluctance of the core and \( R_g \) of the gap between the movable and fixed parts of the core. \( R_c \) is calculated assuming that it has a uniform permeability \( \mu_c \),

\[ R_c = \frac{1}{\mu_c} \int \frac{dl_c}{A(l_c)} = \frac{L_c}{\mu_c A_c}, \]

where the constant \( A_c \) is the harmonic mean of the core section across its length \( L_c \).

To account for magnetic saturation, the relation between the magnetic flux density and the magnetic field in the core, \( B_c \) and \( H_c \), is modeled using a modified version of the Fröhlich-Kennelly equation, as explained in [10],

\[ B_c = \frac{c_1 H_c}{1 + c_2 |H_c|}, \]

where \( c_1 \) and \( c_2 \) are constants to be fitted. From the definition of the magnetic permeability, an expression that relates it to the magnetic flux can be derived,

\[ \mu_c = \frac{B_c}{H_c} = c_1 - c_2 |\phi|. \]

Then, substituting (7) into (5), an expression relating the core reluctance and the magnetic flux is obtained,

\[ R_c = \frac{k_1}{1 - k_2 |\phi|}, \]

where

\[ k_1 = \frac{L_c}{A_c c_1}, \quad k_2 = \frac{c_2}{A_c c_1}. \]

Therefore, the core reluctance can be characterized fitting experimentally the parameters \( k_1 \) and \( k_2 \). On the other hand, a precise empirical characterization of the gap reluctance would require a parametric model that takes into account complex magnetic phenomena, e.g. fringing and leakage fluxes. For that reason, our approach was to obtain the magnetic reluctance of the gap from FEA simulations for several gap lengths. These data, stored as a lookup table and represented in Fig. 2, are used to obtain \( R_g \) for any gap length \( z \) using linear interpolation.

B. Mechanical subsystem

The motion dynamics of the movable part of the valve is given by Newton’s second law, with the position bounded between 0 and \( z_{max} \). This can be modeled as a hybrid automaton with three states (\( q_1, q_2, q_3 \)), as shown in Fig. 3, where \( m \) is the movable part mass and \( F \) is the total force excluding the normal force in case of contact,

\[ F = F_{mag} + F_{sp} + F_f. \]

\( F_{mag}, F_{sp} \) and \( F_f \) are the magnetic, spring and friction forces respectively. Firstly, the magnetic force is calculated...
as a function of $\phi$ and the partial derivative of the reluctance with respect to $z$ \[11\],
\[
F_{\text{mag}} = -\frac{1}{2} \frac{\partial R}{\partial z} \phi^2 = -\frac{1}{2} \frac{\partial R}{\partial z} \phi^2,
\] (11)
where the gap reluctance derivative $\partial R_g/\partial z$ is calculated by linear interpolation of a lookup table, obtained directly from the FEA data of the gap reluctance (Fig. 2). Secondly, the force caused by the compression of the spring is given by Hooke’s law,
\[
F_{sp} = k_{sp}(z_{sp} - z),
\] (12)
where $k_{sp}$ is the stiffness constant of the spring and $z_{sp}$ is the gap length required to leave the spring undeformed. Lastly, the friction is modeled with a viscous force,
\[
F_f = -c_f \dot{z},
\] (13)
where $c_f$ is the damping coefficient.

C. State-space representation

The dynamics of the complete system can be expressed as a set of three nonlinear differential equations: the dynamical equation of the magnetic flux is
\[
\frac{d\phi}{dt} = \frac{v}{N} - \frac{R\phi}{N^2} \left(R_g(z) + \frac{k_1}{1 - k_2|\phi|}\right),
\] (14)
whereas the dynamical equations of the position and velocity are dependent of the discrete state of the hybrid automaton (see Fig. 3). This set of equations can be expressed as a state-space model,
\[
\dot{x} = f(x, q, v, \theta),
\] (15)
where $f$ is the state function, which defines the dynamics of the state vector $x = [\phi \ z \ \dot{z}]^T$ and depends on the state vector itself, the discrete state $q$, the input voltage $v$ and the parameter vector $\theta$,
\[
\theta = [R \ N \ k_1 \ k_2 \ k_{sp} \ z_{sp} \ c_f \ m \ z_{max}]^T.\] (16)

The system output is defined in order to identify $\theta$ by means of comparing the simulated output data with the experimental ones. It is recommended to set the electrical current through the coil $i$ as the system output $y$, because it can be easily measured and can be calculated from the state variables,
\[
y = i = \left(R_g(z) + \frac{k_1}{1 - k_2|\phi|}\right) \phi = \frac{1}{N}. \] (17)

As seen in (17), the observation function $h$ depends on the state and parameter vectors,
\[
y = h(x, \theta).\] (18)

III. MODEL IDENTIFICATION

The model depends on 9 parameters, as stated in (16), that have to be identified. $z_{max}$ and $z_{sp}$ are easily determined by direct measurement of the valve and the spring. To identify the rest, six tests have been performed with a valve, supplying different square wave voltages (Fig. 4) and measuring the resulting currents (Fig. 5).

A. Resistance identification and flux linkage estimation

The flux linkage $\lambda$ is defined as
\[
\lambda = N \dot{\phi}.\] (19)

Although neither $\phi$ nor $\lambda$ are directly measured, $\lambda$ can be estimated offline for identification purposes from the voltage and current measurements if those signals are periodic. An expression of the flux linkage in continuous form can be obtained substituting (19) into (1) and integrating,
\[
\lambda(t_2) = \lambda(t_1) + \int_{t_1}^{t_2} \left(v(t) - R_i(t)\right)dt, \] (20)
and, then, in discrete form with a sampling period $\Delta t$,
\[
\lambda_{j+\delta} \approx \lambda_j + \Delta t \sum_{k=j+1}^{j+\delta} (v_k - R_i_k).\] (21)

If $R$ was known, the flux linkage could be estimated. However, even an apparently negligible error in $R$ would produce cumulating errors in $\lambda$, making it to diverge. To solve this, $R$ is calculated to ensure precisely the periodicity of $\lambda$. The periodic signal repeats itself every $\delta$ samples,
\[
\lambda_{j+\delta} = \lambda_j,\] (22)
so $R$ can be identified substituting (22) into (21), for any $j$,
\[
R = \frac{\sum_{k=j+1}^{j+\delta} v_k}{\sum_{k=j+1}^{j+\delta} i_k}.\] (23)

With the resistance identified, it is possible to estimate the flux linkage between the samples $j$ and $j+\delta$ using (21). Due to measurement and discretization errors, the $R$ values fitted in different periods are slightly different. In addition, $R$ increases with the temperature. To avoid cumulative errors and divergence in the $\lambda$ signal, it is recommended to calculate the value of $R$ in every period.

In Fig. 6, the flux linkage of each one of the tests is represented. The flux linkage estimation can be used for parameter fitting the same way as the measured current.
A. Discretized state-space representation

From this point forward, we distinguish between the voltage and current values $v_k$ and $i_k$ from the noisy measured ones $y_{v,k}$ and $y_{i,k}$. Although the voltage is the input variable, it is indirectly controlled by an electronic circuit, so its value cannot be perfectly known. Therefore it is measured along with the current.

We propose a discrete-time model to be used in the state estimation with discrete-time measurements, taking into account the measurement and discretization errors. Firstly, the magnetic flux equation is obtained applying the Euler method to the differential equation (1). This way, $\phi_k$ depends on the previous values of the flux, current, and voltage,

$$\phi_k = \phi_{k-1} + \frac{\Delta t}{N} (y_{v,k-1} + Ri_{k-1}^2) + w_{\phi,k-1}. \quad (26)$$

Note that, since the true voltage is not accessible, the equation uses the experimental value $y_{v,k-1}$. Consequently, $w_{\phi,k-1}$ accounts for both discretization and voltage measurement errors. The current $i_k$ is modeled as indicated in the expression (17) adding the effects of the error $\varepsilon_{i,k}$ due to model uncertainties,

$$i_k = \left( R_g (z_k) + \frac{k_1}{1-k_2 |\phi_k|} \right) \phi_k - \varepsilon_{i,k}. \quad (27)$$

The magnetic flux difference equation is obtained substituting (27) into (26). Secondly, the position difference equation is defined. During contact, the instantaneous velocity changes abruptly to zero. Due to uncertainty in the previous position $z_{k-1}$, the UKF would estimate a velocity with a large uncertainty variance and, thus, the position $z_k$ would have also a large uncertainty. To work around this issue, the equation of $z_k$ is made dependent on $z_{k-1}$ and $z_{k-2}$ based on the Verlet integration [14].

$$z_k^* = 2z_{k-1} - z_{k-2} + \Delta t^2 \frac{F_{mag} + F_{sp} + F_f}{m} + w_{z,k-1}. \quad (28)$$

where $w_{z,k-1}$ accounts for the discretization error and model uncertainty, and $z_k^*$ is the unbounded position. The position $z_k$ is bounded,

$$z_k = \begin{cases} 0, & z_k^* < 0 \\ z_{max}, & z_k^* > z_{max} \\ z_k^*, & \text{otherwise,} \end{cases} \quad (29)$$

because the actual position cannot surpass the physical limits. $F_{mag}$ and $F_{sp}$ are calculated with (11) and (12) using $\phi_{k-1}$ and $z_{k-1}$. The force $F_f$ depends on the velocity, which can be approximated as the previous average velocity,

$$F_f = -c_f \frac{z_{k-1} - z_{k-2}}{\Delta t}. \quad (30)$$

Lastly, modeling the current error $\varepsilon_{i,k}$ as a white noise is not recommended, as this error has a certain dynamic behavior and is not expected to change abruptly. The dynamics of this error is represented with an autoregressive model, where the error depends linearly on its own previous value,

$$\varepsilon_{i,k} = \varphi \varepsilon_{i,k-1} + w_{i,k-1}. \quad (31)$$
where $\varphi$ is a constant parameter and $w_{i,k-1}$ is white noise.

The state transition equation, formed by (26), (29) and (31), is defined as
\[ x_k = f_i(x_{k-1}, y_{v,k-1}) + w_{k-1}, \]
where $f_i$ is the discretized state transition function, $x_k$ is the new state space vector, which includes the current error $\varepsilon_{i,k}$ as a state variable,
\[ x_k = [\varphi_k \ z_k \ z_{k-1} \ \varepsilon_{i,k}]^T, \]
and $w_k$ is the process errors,
\[ w_k = [w_{\phi,k} \ w_{z,k} \ 0 \ w_{i,k}]^T, \]
which, for estimation purposes, each one is assumed to be Gaussian white noise with zero mean.

The output $y_{i,k}$ is defined as the measured current, that is, the real current $i_k$ (27) plus the measurement error $e_k$,
\[ y_{i,k} = i_k + e_k = \left( R_g(z_k) + \frac{k_i}{1-k_i(\varphi_k)} \right) \frac{\phi_k}{N} - \varepsilon_{i,k} + e_k. \]
Thus, the observation function $h$ of the output $y_k$ is redefined as
\[ y_k = h(x_k, e_k). \]

The identification of the error variances is straightforward. The measurement errors depend on the measurement system. The discretization errors are obtained calculating the discretized variables from the simulated ones and subtracting. The added state variable $\varepsilon_{i,k}$ is calculated subtracting the simulated current from the measured one and, from it, the parameter $\varphi$ and the variance of $w_{i,k-1}$ are obtained.

**B. Algorithms**

The estimation process (Algorithm 1) consists in updating the estimated state $\hat{x}_{k|k}$ and its covariance $P_{k|k}$ in each iteration from previous ones $\hat{x}_{k-1|k-1}$ and $P_{k-1|k-1}$.

**Algorithm 1 Main**

1. Initialize: $\hat{x}_{0|0}, P_{0|0}$
2. Measure: $y_{v,0}, y_{i,0} \leftarrow v(t), i(t)$
3. for $k \leftarrow 1$ to num. iterations do
   4. $\left( \hat{x}_{k|k-1}, P_{k|k-1} \right) \leftarrow $ PREDICTION($\hat{x}_{k-1|k-1}, P_{k-1|k-1}, y_{v,k-1}$)
   5. Measure: $y_{v,k}, y_{i,k} \leftarrow v(t), i(t)$
   6. $\left( \hat{x}_{k|k}, P_{k|k} \right) \leftarrow $ CORRECTION($\hat{x}_{k|k-1}, P_{k|k-1}, y_{i,k}$)
4. end for

First, as described in Algorithm 2, it predicts the next state $\hat{x}_{k|k-1}$ and covariance $P_{k|k-1}$. The UKF method consists in obtaining a set of sigma points $X_{k-1|k-1}$ around the expected estimated state $\hat{x}_{k-1|k-1}$, applying the function $f_i$ to each one to obtain the predicted next states $X_{k|k-1}$, and calculating from it the mean and covariance. In Algorithm 3, the predicted output sigma points $Y_{k|k-1}$ are calculated applying the observation function $h$ to $X_{k|k-1}$. From the

**Algorithm 2 Prediction**

1. function PREDICTION($\hat{x}_{k-1|k-1}, P_{k-1|k-1}, y_{v,k-1}$)
2. $X_{k-1|k-1} \leftarrow $ SIGMA POINTS($\hat{x}_{k-1|k-1}, P_{k-1|k-1}$)
3. $X_{k|k-1} \leftarrow f_i (X_{k-1|k-1}; y_{v,k-1})$
4. $\hat{x}_{k|k-1} \leftarrow $ MEAN($X_{k|k-1}$)
5. $P_{k|k-1} \leftarrow $ COVARIANCE($X_{k|k-1}$)
6. return $X_{k|k-1}, P_{k|k-1}$
7. end function

**Algorithm 3 Correction**

1. function CORRECTION($\hat{x}_{k|k-1}, P_{k|k-1}, X_{k|k-1}, y_{i,k}$)
2. $Y_{k|k-1} \leftarrow h(X_{k|k-1})$
3. $Y_{k|k-1} \leftarrow $ MEAN($Y_{k|k-1}$)
4. $P_{y_k y_k} \leftarrow $ COVARIANCE($Y_{k|k-1}$)
5. $P_{x_k y_k} \leftarrow $ CROSS-COVARIANCE($X_{k|k-1}, Y_{k|k-1}$)
6. $K_{k} \leftarrow P_{x_k y_k} P_{y_k y_k}^{-1}$ ▷ Kalman gain
7. $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_{k}(y_{i,k} - \hat{y}_k)$ ▷ Estimated state
8. $P_{k|k} = P_{k|k-1} + K_{k} P_{y_k y_k} K_{k}^T$ ▷ Error covariance
9. return $X_{k|k}, P_{k|k}$
10. end function

The functions SIGMA POINTS, MEAN, COVARIANCE and CROSS-COVARIANCE are generic functions of the UKF algorithm [9].

**C. Analysis**

The real system measurements cannot be used directly to validate the estimator because the position is not measured. Instead, the input voltage $v$ and output current $i$ data of the model simulations are utilized for validation. The observer is based on the model so the estimations and the simulations are of course expected to match. However, it is possible to artificially add noise and errors to the simulation signals and model parameters, analyzing the robustness of the estimation method to those.

To ascertain the robustness of the UKF estimation to measurement errors, Gaussian white noises with standard deviations $\sigma_v$ and $\sigma_i$, proportional to a normalized standard deviation $\sigma_n$, have been added to $v$ and $i$,
\[ \sigma_v = \bar{v}\sigma_n, \quad \sigma_i = \bar{i}\sigma_n, \]
where $\bar{v}$ and $\bar{i}$ are the mean values of the signals. The UKF algorithm, iterating with a frequency of 50kHz, is capable of estimating the position with considerable accuracy even with large measurement noise (Fig. 8).

To check the robustness of the observer to model inaccuracies, the next analysis is performed comparing the simulation using the identified parameters with estimations based on
The model is nonlinear, and discontinuous when the position of the plunger of a linear travel solenoid valve.

Due to the velocity discontinuity. Fig. 10 shows the estimated position reaches zero with respect to the actual position estimated with real voltage and current data. Due to inaccurate ones. The algorithm is tested applying different relative absolute errors to the six parameters identified in subsection III-B. Each parameter error can be positive or negative, which means that there are $2^6$ possible combinations for each relative error. Testing all combinations for different relative errors, in Fig. 9 we show the worst-case scenarios for each relative error. The estimations are accurate enough with parameter errors around 1%.

Finally, a test is performed with real measurements. Although there are no direct measurements of the actual position, the closing contact instant can be derived from the measured current signal because its dynamics changes abruptly due to the velocity discontinuity. Fig. 10 shows the position estimated with real voltage and current data. Due to model inaccuracies, there is a short delay in the instant that the estimated position reaches zero with respect to the actual contact instant. In case it is needed to reduce the estimation errors, a more complex system model should be used, e.g. taking into account the magnetic hysteresis phenomenon.

V. CONCLUSION

In this paper we have described an observer based on the UKF, with state boundary conditions, to predict online the position of the plunger of a linear travel solenoid valve. The model is nonlinear, and discontinuous when the position reaches one of its physical limits; nevertheless, we have shown that the observer is capable of estimating the position with notable accuracy even with highly noisy measurements and model errors. The test results are obtained using directly the identified parameters and estimated variances. However, for real implementations, adjusting the parameters is recommended to improve the estimation accuracy. The proposed algorithm is specifically designed for controlling a particular solenoid valve. However, by modifying the model, the estimation can be applied for other electromechanical devices in which the position sensing is not affordable or viable, thus increasing the scope of possible applications.

REFERENCES