Diagnostic Signal Method for Fault Identification of Electro-Hydraulic Servo Actuators

Zihan Liu, Prashant N. Kambali, and C. Nataraj

Villanova Center for Analytics of Dynamic Systems, Villanova University, Villanova, PA 19085, USA, (zliu5, prashant.kambali, nataraj)@villanova.edu

ABSTRACT

In this paper, we develop a fault identification approach for electro-hydraulic servo actuators based on injecting a predefined diagnostic signal into the system and then extracting fault-related features from the phase space topology. Next, we build regression models using an artificial neural network, which maps the feature space to fault space to identify the faults represented by the system’s parameters. The performance of the proposed fault identification approach is evaluated when the degradation of permanent armature occurs. The effect of parametric faults on the dynamics is studied and discussed. The robustness of the proposed method under the condition of noise is explored. The obtained results indicate the effectiveness of injected diagnostic signals in enriching the dynamics of the system and increasing the quality of extracted for training artificial neural networks.

1. INTRODUCTION

Electro-hydraulic servo systems perform precise and sensitive tasks such as controlling the position or force, changing the pressure, and starting or stopping the hydraulic flow (Tamburrano, Plummer, Distaso, & Amirante, 2018). A hydraulic system is vulnerable to dust particles, erosion, and unstable working conditions such as a high load or severe impact which lead to unpredictable parametric changes in friction, supply pressure, or fluid properties. If these initial abnormalities are not identified in time, they may result in the deterioration of performance even failure or instability of the whole system (Ren, Chen, Hu, & Yuan, 2016). Therefore, developing an effective diagnostic technique for continuously monitoring the condition and identifying the fault parameters at their very inception is of high importance.

Machinery faults can be identified either by interrupting the machine to run a diagnostic procedure or by performing an online diagnosis during the normal operation of the machine.

In most cases, the latter has evident advantages in safety and economy. Additionally, because of the uncertain hydraulic parameters caused by viscous damping, variable load stiffness, and external loads (Guo & Chen, 2021), it is usually necessary to develop a tracking controller to ensure asymptotic stability and adaptivity of hydraulic systems (Z. Yao, Yao, & Sun, 2018; Deng & Yao, 2019). Therefore, the development of online diagnostic and online state monitoring techniques which play an important role in controlling is of great significance in hydraulic systems.

Online diagnosis targets incipient faults which are harmless to the normal operation of systems. Because of this, the variations of state variables caused by parametric faults are barely detectable. One feasible solution is deriving new healthy indicators as a function of fault parameters which are more sensitive to parametric faults according to the physics model of systems (Ersfolk et al., 2018; Soualhi, Nguyen, Medjaher, Lebel, & Cazaban, 2022; G.-W. Kim & Wang, 2007; Lian, Xu, & Lu, 2013). Then, extra sensors are required to measure new data to obtain healthy indicators for parametric fault identification. However, the requirements of extra hardware and an accurate physics model limit the applicability of the above method.

Another promising method of online diagnostics is the injection of diagnostic signals to increase the sensitivity of state variables to parametric faults. The concept of signal injection was originally developed for rotor position estimation (Jiang & Holtz, 1997). Briz et al. (Briz, Degner, Diez, & Guerrero, 2004) injected high-frequency signals to detect the stator winding faults in inverter-fed machines. Dotelli et al. (C. Yao et al., 2014) used high-frequency signals to monitor membrane drying and low-frequency signals to detect cell flooding in polymer electrolyte membrane fuel cells. Yao et al. (C. Yao et al., 2014) proposed a diagnostic algorithm for online monitoring power transformer winding deformation by injecting high-frequency signals.

However, there are no studies in literature regarding online diagnostics of hydraulic servo systems employing signal injection techniques. Most efforts in hydraulic diagnostics have
been devoted to improving the efficiency of utilizing the fault-related information in the measured signal by means of signal processing (Goharrizi & Sepehri, 2011; Amirat, Choqueuse, & Benbouzid, 2013), feature extraction (Jegadeeshwaran & Sugumaran, 2015; Qiu, Min, Wang, & Fan, 2022), and machine learning algorithm (Tang, Zhu, & Yuan, 2021; Huang, Wu, Li, Yang, & Gui, 2021). Nevertheless, our team still believes in the promising prospects of extending the technique of signal injection to hydraulic systems. Because the fundamental obstacle of online diagnostics is the limited fault-related information resulting from incipient faults. Compared with increasing the utilization efficiency of fault-related information, using injected diagnostic signals to enrich the fault-related information is more potential.

The difficulty in extending the technique of signal injection to hydraulic servo systems results from the nonlinear nature and coupled mechanisms of electro-hydraulic servo systems. The sensitivity to fault parameters cannot be guaranteed by established physics models. Meanwhile, gradual degradation in several mechanisms co-evolves, it is difficult to obtain independent and separable responses excited by injecting diagnostic signals. Therefore, the effectiveness of frequency analysis which is widely used in signal injection techniques is compromised. Moreover, multiple-factorial noises obscure injected diagnostic signal which is designed with low power to ensure normal operation of systems. In this study, we develop a novel fault identification approach that is based on injecting a predefined diagnostic signal into the system and then extracting fault-related features from phase space topology. The injected diagnostic signal is designed to enrich the fundamental response instead of generating a separable response. Meanwhile, the abundant dynamics phenomena observed in the enriched responses ensure the sensitivity to fault parameters. The feature extraction technique is chosen to be the phase space topology (PST) method which is conceived and developed by our own team (Samadani, Kwuimy, & Nataraj, 2015; Mohamad, Nazari, & Nataraj, 2020). The PST method extracts density-based features from the phase space density distributions which circumvent frequency analysis and the influence of noises.

The rest of this paper is organized as follows. In Section 2, a description of servo systems is presented. In Section 3, the proposed diagnostic method is explained. In Section 4, a parametric analysis is performed on the fault’s effects on the system response. Section 5 describes the parametric estimation algorithm and presents the results. Finally, Section 6 summarizes the main conclusions of the paper.

2. Electro-Hydraulic Servo System

An electro-hydraulic servo actuator consists of a symmetrical cylinder controlled by a two-stage servo valve as shown in Fig 1. The following subsections will briefly introduce the system and its faults.

2.1. Working Principle

In general, electro-hydraulic servo actuator systems consist of a torque motor and two stages of hydraulic power regulation. The working principle of the electro-hydraulic servo actuator can be summarized as follows. The electromagnetic torque motor controls the first stage by positioning a flapper. The flapper controls the hydraulic fluid flow through two nozzles in the first stage. This process provides positive internal closed-loop flow control. Next, differential pressure positions the second-stage spool, which controls the direction and rate of hydraulic fluid flow to the actuator.

A nonlinear dynamical model of a two-stage servo actuator system with spool position feedback has been used in the analysis (Gordić, Babić, & Jovičić, 2004).

Figure 1. Functional schematic of an electro-hydraulic servo actuator

2.2. Parametric Fault

In industrial servo systems, one main performance requirement for a successful application is that the electro-hydraulic system has to meticulously achieve all the desired control variables such as position, velocity, and acceleration in addition to the desired force and torque. There are many parameters that can be monitored in typical electro-hydraulic systems. In this study, we focused on the degradation of the permanent magnet of the armature.

The temperature of motors typically rises during high-speed motion, which has a significant effect on the permanent magnet of the motor and the torque capabilities. The motor torque
constant \((K_i)\) is directly related to the magnetic flux density of the permanent magnets. The overall flux density changes with an increase in the magnet temperature, and the increase rate of temperature also varies depending on the magnet material.

If the motor is operating within the designed operating temperatures, the decrease in flux density is temporary and will be recovered as the magnet temperature cools down. This is known as reversible thermal demagnetization. However, if the temperature of the magnets exceeds the maximum designed operating temperature, partial demagnetization will occur and will be permanent, which is also known as irreversible thermal demagnetization.

Reversible thermal demagnetization depends on the specific magnet material that is being used. The reversible decrease in \(K_i(T)\) with increasing magnet temperature is given by the following formula:

\[
K_i(T) = K_i(T_0)[1 - \alpha_T(T - T_0)]
\]

where, \(T_0 = 25^\circ C\) is the initial temperature. Both reversible and irreversible phenomena affect the motor torque constant \(K_i\), and hence, the torque capability and the permanent magnet motor efficiency are dependent.

### 3. Methodology

An overview of the proposed fault identification method used in this paper is provided in Fig. 2. The main idea is based on injecting a pre-defined diagnostic signal into the system and then analyzing the system response represented by its state variables using the phase space topology (PST) method to extract fault-related features. Next, a regression model using an artificial neural network (ANN) is developed to estimate the electric-hydraulic servo actuator parameters. In summary, features will be extracted from the phase space of state variables after injecting a signal into the system. Then, these features will be used for estimating the degradation of the permanent magnet of the valve armature, represented by the change of the coefficient \(K_i\).

#### 3.1. Injection of Diagnostic Signal

For the identification of faults, we propose to inject a pre-defined signal \(s_I\) into the servo-actuator system. A periodic signal was selected as an input as follows.

\[
s_I = A_I \sin(f_I t)
\]

where \(A_I\) and \(f_I\) are the amplitude and the frequency of the injected signal, respectively. The design of the injected signal depends on the dynamic analysis of the electric-hydraulic servo system. The sine function is chosen to maximize the dynamic information in the obtained response, the selection of amplitude ensures the normal and safe operation of the system, and the chosen frequency should circumvent the resonance frequencies.

The original excitation is assumed as a periodic signal as follows.

\[
s_p = A_p \sin(f_p t)
\]

where \(A_p\) and \(f_p\) are the amplitude and the frequency of the periodic input signal, respectively.

The main idea here is to inject a signal at a specific frequency which can enrich the nonlinear response and unfold some hidden information about the system’s health condition. This signal should not affect the usual operation of the system. In fact, using a low amplitude and high frequency injected signal has proven its reliability for fault detection in motor applications (S.-K. Kim & Seok, 2011; Arellano-Padilla, Sumner, & Gerada, 2011). Hence, the amplitude of the original excited signals is assumed as \(A_p = 0.04\) and the injected diagnostic signal amplitude is selected to be a smaller value \(A_I = 0.01\). Because the goal is to develop a diagnostic approach that will not cause harm to the system, the injected diagnostic signal frequency is selected to be \(f_I = 800\) Hz while the periodic input signal frequency is selected to be \(f_p = 500\) Hz.

Figs. 3 demonstrate the time series of angular displacement excited with only periodic input signal \(s_p\) (a), and both periodic input signal \(s_p\) and injected diagnostic signal \(s_I\) (b). Fig 4 demonstrate the power spectra of angular displacement excited with periodic signal \(s_p\) (a) and both periodic signal \(s_p\) and injected diagnostic signal \(s_I\) (b). With the injected diagnostic signal, multiple periods appear in the time series as shown in Fig. 3 (b). And, the multiple peaks in the spectrum
Figure 3. Time series responses of angular displacement $\theta$: (a) with excitation $s_p$, (b) with excitation $s_p$ and injected diagnostic signal $s_I$

Figure 4. Spectrum of angular displacement $\theta$: (a) with excitation $s_p$, (b) with excitation $s_p$ and injected diagnostic signal $s_I$

Further prove the existence of multiple periodic responses as shown in Fig. 4 (b).

Note that, the injected diagnostic signal enriches the dynamics, and these changes remain on a low scale in terms of amplitude which will not sabotage the normal operation of the system. In the next section, we discuss the selection of features that can capture these low-scale changes.

3.2. Phase Space Topology Method

Although phase space is a widespread concept in dynamic system theory, depicting all possible dynamical states of the system evolving over time, we have harnessed it into a new and powerful technique called phase space topology (PST). PST is based on the characterization of the phase space by means of geometry properties, such as eccentricity, harmonic amplitude, excursion variation, and density distribution peak properties or properties derived from them, and has proven to be a very effective approach that does not rely on inherent assumptions of linearity.

There is a different phase space trajectory for each type of system response. For example, a single-frequency periodic response has a single loop closed, whereas multi-periodic and quasi-periodic responses have trajectories of several loops. This implies that the phase space trajectory can be used to describe the nature of the system in a qualitative fashion. The method of PST is, however, based on characterizing the phase space topology with quantitative measures by creating the phase space density distribution for each axis.

A kernel density estimator was used as follows in order to estimate the density distribution. Let $X = (x_1, x_2, \ldots, x_n)$ be a set of independent and identically distributed sample data drawn from an unknown density distribution function $\Psi$. Its kernel density estimator, $\hat{\Psi}_h(x)$, can estimate the shape of this function, where $\sim$ implies that it is an estimate, and $h$ indicates that its value may be dependent on $h$.

\[
\hat{\Psi}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} \Gamma \left( \frac{x - x_i}{h} \right)
\]

where, $h > 0$ is the smoothing parameter, and the kernel function $\Gamma(.)$ satisfies the requirements as follows.

\[
\int_{-\infty}^{\infty} \Gamma(u) du = 1
\]

\[
\Gamma(-u) = \Gamma(u) \forall u
\]

The standard Gaussian density function is used due to its convenient mathematical properties, which is defined as the following equations.

\[
\Gamma(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}
\]

After estimating the kernel density for phase space variables, quantitative measures are extracted from these distributions. For example, the phase space plane for the angular displacement $\theta$ and angular velocity $\dot{\theta}$ of the responses excited with periodic signal $s_p$ is presented in Fig. 5 (a). A single loop can be observed in the phase plane. The corresponding density distribution of the variable $\theta$ is estimated and presented in Fig. 5 (b). As is evident, two unique sharp peaks are caused by the higher concentration of points marked as red dots in the corresponding areas of the phase plane as shown in Fig. 5 (d). These points, along the vertical axis in the phase plot, are called the returning points and are the areas in which the system spends more time. Likewise, the density distribution for the phase plane variable $\dot{\theta}$ is shown in Fig. 5 (c). Two peaks corresponding to the returning points along the horizontal axis of the phase plane are marked as black dots in Fig. 5 (d) to emphasize these points.

After the injection of the diagnostic signal $s_I$, the changes in the phase space plane are demonstrated in Fig. 6 (a). We can observe multiple loops in the phase plane. The corresponding density distribution of the variable $\theta$ is estimated and presented in Fig. 6 (b). Four unique sharp peaks are caused by the higher concentration of points marked as red dots in the corresponding areas of the phase plane as shown in Fig. 6...
The coefficients of the Legendre polynomials are obtained by periodic excitation $s_p$: (a) phase plane, (b) density distribution of $\theta$, (c) density distribution of $\dot{\theta}$, (d) phase plane with the location of the corresponding distribution peaks

(d). The density distribution for the phase plane variable $\dot{\theta}$ is shown in Fig. 6(c). Six peaks are caused by the returning points along the phase plane marked as black dots in the corresponding areas of the phase plane as shown in Fig. 6(d) to emphasize these points.

It is worth noting that for dimensions higher than three, even though visualization of the phase space trajectory is impossible, the method is still applicable. This is because the computations are performed individually and independently for each state of the system.

The next step includes extracting accurate and adequate information that fully represents the system’s behavior. In order to do that, while preserving as much information as possible, we propose a feature set by approximating the phase space density distributions using Legendre Polynomials. When the approximation matches the actual density distribution, the density distributions using Legendre Polynomials. When the estimation coefficient (PCC) are calculated using the following equations to determine the fit quality.

\[
\text{RMSE} = \sqrt{\frac{1}{N} Z Z^T}, \quad \text{PCC} = \frac{\sigma_a^T \sigma_d}{\sqrt{\sigma_a^T \sigma_d (\sigma_d^T \sigma_a)}}
\]

where, $Z = (y_d - \Psi_a)$ is the residual vector, $N$ is the number of data points, $\sigma_a$ is the actual density, $\sigma_d$ is the density distribution using Legendre Polynomials.

The coefficients $\hat{\beta}$ constitute the features in the PST approach. The approximate density is then calculated using Legendre Polynomials as follows.

\[
\Psi_a = X \hat{\beta}
\]

using the least squares method assuming the following linear regression model as the following form.

\[
\Psi(x, \beta) = \sum_{j=1}^{m} \beta_j \Phi_j(x) \tag{10}
\]

Letting

\[
X_{ij} = \frac{\partial \Psi(x_i, \beta)}{\partial \beta_j} = \Phi_j(x_i), \quad \text{for} \quad i = 0, 1, 2, \ldots \tag{11}
\]

the estimated coefficients are given by the following equation.

\[
\hat{\beta} = (X^T X)^{-1} X^T y_d \tag{12}
\]
of points in the density function, \( \sigma_d = \langle y_d - E\{y_d\} \rangle \) and \( \sigma_a = \langle \Psi_a - E\{\Psi_a\} \rangle \), where \( E\{\cdot\} \) is the expected value.

### 3.3. Artificial Neural Network

Artificial neural networks (ANNs) are computational models that were historically inspired by the biological architecture of the human brain. Essentially, ANNs are statistical data modeling tools that map complex input-output relationships and recognize the behavioral patterns of nonlinear systems. The most popular type of ANN is the Multi-Layer Feed Forward (MLFF) neural network.

An MLFF neural network includes different layers of neurons, including an input layer, hidden layers (which can be one or more layers), and an output layer that is connected with weights. Neurons in all layers, except the input layer, use predefined activation functions. In this study, the structure of ANN has one hidden layer. The activation function of all neurons is \( \text{Tansig} \), which is defined as the following equation.

\[
f_{\text{Tansig}}(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{15}
\]

The inputs of ANN in this study are PST features and the output is the interest parameters \( K_i \). The PST features in this study are the Legendre polynomial coefficients obtained from state variables.

### 4. Analysis and Discussion

#### 4.1. Fault Sensitivity Analysis

The change in dynamic responses of the hydraulic systems can be used to identify the faults. To capture the variation of \( K_i \) from the healthy region, \( C_{K_i} \) is defined in Eq. 16.

\[
C_{K_i} = \frac{K_i}{\overline{K_i}} \tag{16}
\]

\( C_{K_i} \) are the values of \( K_i \) normalized to what we define to be healthy value \( \overline{K_i} \). The system is considered to operate in a healthy regime when \( C_{K_i} \) is close to 1. For illustration, we assume that \( \overline{K_i} = 0.556 \text{ Nm/A} \). Generally, nonlinear systems might have multiple healthy regimes, and hence the healthy values for the same parameter are non-unique. However, it should be emphasized that our approach has no inherent assumptions about the number of healthy regions and what the healthy value of the system might be.

To analyze how the system response is affected by the variation of these parameters, four faulty conditions are considered and compared with the healthy condition. These faulty conditions are caused by the torque gain degradation corresponding to 20\%, 40\%, 60\%, and 80\% decrease in \( C_{K_i} \).

#### 4.2. Noise Effect Analysis

Noise from different sources, i.e., fluid dynamic or mechanical noise (Peng, Li, & Fan, 2014), is likely to contaminate the response measured in reality. Hence, in order to introduce a sense of reality into our own model, we will integrate measurement noise as follows. Signal-to-noise ratio (SNR) describes the intensity of the noise and is expressed as Eq. 17.

\[
\text{SNR} = 20 \log_{10} \left( \frac{\sigma_r}{\sigma_n} \right) \tag{17}
\]

where \( \sigma_r \) and \( \sigma_n \) are the root mean square values of the response signal and the noise, respectively. Small SNR indicates large noises and vice versa.

Figure 7 demonstrates the time series responses of angular displacement \( \theta \) with (a) period signal \( s_p \) and (b) period signal \( s_p \) and injected diagnostic signal \( s_I \).

#### 4.3. Phase Space Density Analysis

The abundant nonlinear dynamics driven by the injected diagnostic signal leads to variations in the phase space density distribution.

In case of the periodic excitation, the density distributions of the angular displacement and velocity without injected diagnostic signal for the various cases of \( C_{K_i} \) are demonstrated in Fig. 8. As \( C_{K_i} \) decreases, we observe the increases in amplitude of density distribution indicated by the \( y \)-axis and the decrease in the spread of density distribution on the \( x \)-axis. The density distributions are clearly distinguishable with the variation of \( C_{K_i} \) for both angular displacement and velocity as shown in Figs. 8 (a) and (b).

Fig. 9 demonstrates the density distributions of angular displacement and velocity with combined periodic signal and injected diagnostic signal. It can be observed that more peaks appear as \( C_{K_i} \) increases and the peaks are clearer even for small values of \( C_{K_i} \), as compared with density distribution without injected signal as shown in Fig. 8.
From the above, we conclude that the injection of the diagnostic signal enriches the dynamics of the hydraulic-electro servo system and provides more qualitative characteristics in the phase space plane and density plots which can be extracted as features for developing a machine-learning model.

5. Parameter Estimation and Results

In this section, 100 sets of data corresponding to various $C_{I_0}$ are generated as the training set to develop a neural network. Other 20 sets of data are generated as the testing set.

In the case of periodic excitation, PST features are extracted from both angular displacement $\theta$ and angular velocity $\dot{\theta}$ separately to train NNs, and the estimated regression plots of them are demonstrated in Fig. 10. Fig. 10 (a) shows the outputs of NNs trained by $\dot{\theta}$-related features where the RMSE of the case without the injected diagnostic signal is $8.2157e-04$, and the RMSE of the case with the injected diagnostic signal is $2.3506e-05$. Fig. 10 (b) demonstrates the outputs of NNs trained by $\theta$-related features where the RMSE of the case without the injected diagnostic signal is $1.8965e-04$, and the RMSE of the case with the injected diagnostic signal is $1.4795e-05$.

After the introduction of noise with $SNR = 5$ dB in the simulations, the corresponding regression plots are demonstrated in Fig. 11. Fig. 11 (a) shows the outputs of NNs trained by $\theta$-related features where the RMSE of the case without the injected diagnostic signal is $0.0249$, and the RMSE of the case with the injected diagnostic signal is $0.0039$. Fig. 11 (b) exhibits the outputs of NNs trained by $\dot{\theta}$-related features where the RMSE of the case without the injected diagnostic signal is $0.0373$, and the RMSE of the case with the injected diagnostic signal is $0.0073$.

Note that the performance of NN’s is better using the features of density distribution plots excited with combined periodic signal $s_p$ and injected diagnostic signal $s_I$. This can be observed by the RMSE values tabulated in Table 1 which validates the effectiveness of the proposed method in enhancing the robustness of extracted features in case of added noise. Moreover, we can observe lower RMSEs with injected diagnostic signals for each scenario in Table 1. We conclude that the injected diagnostic signal is effective in increasing the estimation accuracy of trained NNs.

6. Conclusion

In this study, we propose an online diagnostic method using injected diagnostic signal to identify the degradation of
permanent armature represented by the motor torque constant gain ($K_t$) in an electro-hydraulic servo system. The injected diagnostic signal enriches the dynamics of the electrohydraulic system which leads to more distinguishable density distribution plots of the phase portraits. Therefore, the features extracted from phase space by PST methods are more capable of capturing the fault-related information and training an artificial neural network for parametric fault estimation. The lower RMSEs of estimating fault parameters using the extracted features from the PST method for each scenario with injected diagnostic signal demonstrate the effectiveness of the proposed method. In case of added noise, the artificial neural network performance better estimations with injected diagnostic signal compared to without injected diagnostic signal with lower RMSE values. The presented results demonstrate the robustness of the proposed fault diagnostic method for the hydraulic system.

the evident superiority of RMSEs with injected diagnostic signals compared with RMSEs without injected diagnostic also validate the effectiveness of the proposed method in enhancing the robustness of trained ANNs.

**ACKNOWLEDGMENT**

This project was supported by research grants from the Office of Naval Research (Grant No. N00014-19-1-2070 and N00014-22-1-2480). We are grateful to ONR and Capt. Lynn Petersen, the program manager for recognizing the importance of this work.

**REFERENCES**


