Abstract—In this paper, a new type of interacting multiple model (IMM) is introduced for the purposes of fault detection and diagnosis. The standard IMM is combined with a relatively new filtering method referred to as the smooth variable structure filter (SVSF). The SVSF is a type of sliding mode estimator, formulated in a predictor-correct fashion. It keeps the estimated state close to the true trajectory, and creates a stable estimation process. The combined method, referred to as the SVSF-IMM, is applied to an electrohydrostatic actuator (EHA). The results of the experiment are compared with the common form of the IMM, which utilizes the popular Kalman filter (KF).

I. INTRODUCTION

The ability to detect and diagnose faults is essential for the safe and reliable control of mechanical and electrical systems. In the presence of a fault, the system behaviour may become unpredictable, resulting in a loss of control which can cause unwanted downtime as well as damage to the system. There are two main types of methods to detect and diagnose faults: signal-based and model-based [1]. Signal-based fault detection methods typically use thresholds to extract information from available measurements [2,3]. This information is then used to determine if a fault is present. Model-based methods, as the name suggests, makes use of faults which can be modeled, typically through system identification. This type of fault detection and diagnosis is popular when well-defined models can be created and utilized.

The interacting multiple model (IMM) strategy may be utilized for fault detection and diagnosis, and is classified as a model-based method [4]. Essentially, the IMM makes use of a number of models (for each fault condition), and is associated with filters that run in parallel. The output from each filter includes the state estimate, the covariance, and the likelihood calculation (which is a function of the measurement error and innovation covariance). The output from the filters is used to calculate fault probabilities, which gives an indication of how close the filtered model is to the true fault model. The IMM has been shown to work significantly better than single model methods, since it is able to make use of more information [5]. It also works extremely well for standard estimation problems such as target tracking, where there are typically two models used (i.e., uniform motion or coordinated turn) [5].

The standard IMM employs the use of the Kalman filter (KF), which is a very popular and well-studied estimation method. Introduced in the early 1960’s, it yields a statistically optimal solution for linear estimation problems in the presence of Gaussian noise [6]. In other words, based on the available information on the system, it yields the best possible solution in terms of estimation error [7]. The KF is formulated in a predictor-corrector manner, such that one first predicts the state estimates using knowledge of the system model. These estimates are termed as a priori, meaning ‘prior to’ knowledge of the observations. A correction term is then added based on the innovation (also called residuals or measurement errors), thus forming the updated or a posteriori (meaning ‘subsequent to’ the observations) state estimates [8].

The KF assumes that the system model is known and is linear, the system and measurement noises are white, and the states have initial conditions and are modeled as random variables with known means and variances [5,9]. However, these assumptions do not always hold in real applications. If one of these assumptions is violated, the KF performance becomes sub-optimal and could potentially become unstable [10]. Moreover, the KF is sensitive to the machines arithmetic precision and the complexity of the calculation (in particular, the inversion operator). The smooth variable structure filter (SVSF) was introduced in an effort to provide a more stable filter, while maintaining a relatively good estimate [11,12]. The SVSF is a type of sliding mode estimator, where gain switching is used to ensure that the estimates converge to within a boundary of the true state values (i.e., existence subspace) [13]. In its present form, the SVSF is stable and robust to modeling uncertainties and noise, given an upper bound on the level of un-modeled dynamics or knowledge of the magnitude of noise. It has been shown to work very well when the system is not well-defined or there are modeling errors.

In an effort to improve on the standard IMM, this paper combines the SVSF with the IMM, thus creating a new fault detection and diagnosis method referred to as the SVSF-IMM. The following section introduces the SVSF and the formulation of the SVSF-IMM. The method is then applied on a mechanical system to demonstrate its effectiveness.
II. THE IMM-SVSF STRATEGY

A. Smooth Variable Structure Filter

The basic estimation concept of the SVSF is shown in the following figure. Some initial values of the estimated states are made based on probability distributions or designer knowledge. An area around the true system state trajectory is defined as the existence subspace. Through the use of the SVSF gain, the estimated state will be forced to within this region. Once the value enters the existence subspace, the estimated state is forced into switching along the system state trajectory. A saturation term may be used in this region to reduce the magnitude of chattering or smooth-out the result. As previously mentioned, the SVSF gain introduces a certain amount of chattering which brings an inherent amount of stability [14]. This makes the estimation strategy an attractive method for control problems when not all of the dynamics are well known or defined correctly.

Fig. 1. The smooth variable structure filter estimation concept is shown in the above figure [12].

The SVSF method is model based and applies to smooth linear or nonlinear dynamic equations. The estimation process is iterative and may be summarized by the following set of equations (for a linear control or estimation problem). Like the KF, the system model is used to calculate \( a \) \( \textit{priori} \) state and measurement estimates. A corrective term, referred to as the SVSF gain, is calculated as a function of the error in the predicted output and a smoothing boundary layer. This gain is then used to update the state estimates. Note that the estimation process is stable due to the gain calculation of (4), which keeps the estimates bounded. Refer to the Appendix for a list of pertinent nomenclature and variable definitions.

\[
\hat{x}_{k+1|k} = \hat{A} \hat{x}_{k|k} + B u_k \\
\hat{z}_{k+1|k} = \hat{H} \hat{x}_{k+1|k} \\
e_{z_{k+1|k}} = z_{k+1} - \hat{z}_{k+1|k} \\
K_{k+1} = \left( e_{z_{k+1|k}}^T + \gamma \left| e_{z_{k|k}} \right| \right) \cdot \text{sat} \left( \frac{e_{z_{k+1|k}}}{\psi} \right) \\
\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1} e_{z_{k+1|k}} \\
e_{z_{k+1|k+1}} = z_{k+1} - \hat{H} \hat{x}_{k+1|k+1}
\]

Furthermore, the switching found within the existence subspace is smoothed out by using the saturation term of (4), which is defined by the \( a \) \( \textit{priori} \) output error and some predetermined boundary layer width. In its current form, the boundary layer width requires tuning by trial-and-error, based on some designer knowledge of the uncertainties and noise levels. Two critical variables in the SVSF estimation process are the \( a \) \( \textit{priori} \) and the \( a \) \( \textit{posteriori} \) output error estimates, defined by (3) and (6), respectively.

B. Formulation of the IMM-SVSF

The SVSF provides an estimation process that is sub-optimal albeit stable. Therefore, utilizing a multiple model strategy which increases the overall accuracy of the estimation process is beneficial. However, it is important to note that the SVSF in the form presented in [12] could not be integrated with the IMM strategy. The standard IMM utilizes covariance outputs from the filter, which the previous form of the SVSF did not include. In [15], an augmented form of the SVSF with a covariance calculation was presented, and will be used to combine the SVSF with the IMM strategy.

\[
\begin{align*}
\hat{x}^1_{k+1|k+1} & = \hat{x}^1_{k+1|k} + K_{k+1} e_{z_{k+1|k}} \\
\hat{x}^2_{k+1|k+1} & = \hat{x}^2_{k+1|k} + K_{k+1} e_{z_{k+1|k}} \\
A'_{k+1} & = \text{Mode Probability Update and Mixing Probability Calcs.} \\
A''_{k+1} & = \text{State Estimate and Covariance Combination}
\end{align*}
\]

Fig. 2. The interacting multiple model (IMM) strategy adapted from [5] is shown in the above figure.

The IMM was implemented as per Section 11.6 of [5]. The concept is shown in the above figure. Essentially, any estimation strategy (with a covariance derivation) may be applied on the models of interest. In this case there are two models. Prior to feeding the initial estimates and covariance’s into the filter models, an interaction (mixing) stage takes place, as per the following equations.

\[
\mu_{ij|k} = \frac{1}{C_j} p_{ij} \mu_{ik} \\
c_j = \sum_{i=1}^{n} p_{ij} \mu_{ik}
\]

1397
$$\hat{x}_{k|k}^0 = \sum_{i=1}^{r} \mu_{i|k} \hat{x}_{i|k}^{0j}$$  \hspace{1cm} (9)$$

$$p_{k|j}^0 = \sum_{i=1}^{r} \mu_{i|k} \left( p_{i|k}^{j} + \left[ \hat{x}_{i|k}^{0j} - \hat{x}_{k|k}^{0j} \right] \left[ \hat{x}_{i|k}^{0j} - \hat{x}_{k|k}^{0j} \right]^T \right)$$ \hspace{1cm} (10)

The predicted mode probability \( \mu \), as defined by (7), is first calculated based on the mode transition matrix (user defined) and the previous (or initial) mode probabilities. The mode probabilities are used like weights to determine the corresponding initial estimates and covariance. Using these values and the measurement as inputs, the SVSF calculates the corresponding estimates (16) and covariance (17), as follows:

$$\hat{x}_{k+1|k}^i = \hat{A}_{k|k}^{i} + B_{k}^{i} u_k$$  \hspace{1cm} (11)$$

$$p_{k+1|k}^{i} = \hat{A}_{k|k}^{i} p_{k|k}^{i} \hat{A}_{k+1|k}^{iT} + Q_k$$  \hspace{1cm} (12)$$

$$s_{k+1|k}^i = \hat{H}_{k+1}^{i} p_{k+1|k}^{i} \hat{H}_{k+1|k}^{iT} + R_{k+1}$$  \hspace{1cm} (13)$$

$$e_{k+1|k}^j = z_{k+1} - \hat{H} \hat{x}_{k+1|k}^j$$  \hspace{1cm} (14)$$

$$K_{k+1}^j = \text{diag} \left( \left[ e_{k+1|k}^j + \gamma \frac{e_{k+1|k}^j}{\sqrt{\text{diag} \left( e_{k+1|k}^j \right)}} \right] \right)^{-1}$$  \hspace{1cm} (15)$$

Based on the innovation matrix (13) and the a priori measurement error (14), a likelihood function is calculated, as follows:

$$L_{k+1}^j = \frac{1}{\sqrt{2\pi S_{k+1|k}^{j}}} \exp \left( -0.5 \frac{e_{k+1|k}^j e_{k+1|k}^j}{S_{k+1|k}^{j}} \right)$$  \hspace{1cm} (19)$$

These likelihood functions \( L \) are used to determine updates to the mode probability and mixing calculations, as follows:

$$\mu_{j|k+1} = \frac{1}{c_j} L_{k+1}^j \tilde{e}_j$$  \hspace{1cm} (20)$$

$$c_j = \sum_{j=1}^{r} L_{k+1}^j \tilde{e}_j$$  \hspace{1cm} (21)$$

Furthermore, the two sets of state estimates and covariance’s may be combined (for output purposes only) to determine the IMM estimate of the respective filter (i.e., KF or SVSF), as per (22) and (23) [5]:

$$\hat{x}_{k+1|k+1} = \sum_{j=1}^{r} \mu_{j|k+1} \hat{x}_{k+1|k+1}^j$$  \hspace{1cm} (22)$$

$$p_{k+1|k+1} = \sum_{j=1}^{r} \mu_{j|k+1} \left( p_{k+1|k+1}^j + \left[ \hat{x}_{k+1|k+1}^j - \hat{x}_{k+1|k+1} \right] \right)$$  \hspace{1cm} (23)$$

The main IMM-SVSF process and equations are defined by (7) through (23), where there are \( i, j = 1, ..., r \) models.

III. EXPERIMENTAL SETUP

The experimental setup used in this paper involved an electrohydraulic actuator (EHA). An EHA is an emerging type of actuator typically used in the aerospace industry, and are self-contained units comprised of their own pump, hydraulic circuit, and actuating cylinder [16]. The main components of an EHA include a variable speed motor, an external gear pump, an accumulator, inner circuit check valves, a double-rod double-acting cylinder, and a bi-directional pressure relief mechanism. The schematic of the EHA circuitry is shown in the following figure.

Fig. 3. A typical electrohydrostatic actuator (EHA) circuit diagram is shown above [16]. The EHA used in the experiment did not contain the filter sub-circuits.

The EHA can be divided into two subsystems. The first is the inner circuit that includes the accumulator and its surrounding check valves. The second is the high pressure outer circuit which performs the actuation. The inner circuit prevents cavitation which occurs when the inlet pressure reaches near vacuum pressures and provides make-up fluid for any dynamic leakage [16]. This section is statically charged to 276KPa (40psi) which is enough pressure to avoid cavitation but it is also low enough to allow flow from the case drain back into the circuit. The inner circuit during normal operation is negligible in mathematical modeling. Mathematical modeling of the EHA has been performed and can be seen in detail in [16].

A dual version of the EHA was developed that places two systems in series by rigidly attaching the shafts of both cylinders to one another. This system also includes two 2-way, normally closed solenoid valves that act as bypass valves in the event of a fault. This allows one pump to drive
both cylinders if needed. There are also valves used to connect the inlet and outlet lines of both axes to each other. This provides a steady motion if the actuation of both pumps are not equal. The inclusion of the throttling valves also allows for fault simulations. A figure of the experimental setup used in this paper is shown below.

The two faults that will be introduced to this system are increased friction and internal leakage. To incur these faults, one of the axes will be used as the driving mechanism while the other will act as a load. Refer to the figure in the Appendix for a more detailed circuit diagram of the EHA. Axis A pump is used to drive both cylinders while the axis B blocking valve is in the closed position for all scenarios. The axis B blocking valve is closed, which decouples the axis B gear pump from its corresponding cylinder. In this case, the valve that allows fluid to flow between both chambers of the axis B cylinder is the axis B throttle valve (axis B throttle blocking valve is open). This valve, along with its counterpart on axis A, is a normally open, 2-way, bidirectional proportional valve that receives a 0-10V input from the controller. The increased throttling of this valve while the cylinder is in motion increases the back pressure, which simulates increased friction in the system. To simulate internal leakage, axis A throttling valve is used (axis A throttle blocking valve is open). As the axis A gear pump is in rotation, the axis A throttling valve incurs cross-port leakage between both chambers of its corresponding cylinder.

Three operating modes were tested in this experiment: normal mode, friction mode, and leakage mode. For each operating mode a dynamic model was obtained using black-box system identification methods. The operating range used for the servo-motor is ±1.5V which corresponds to a motor speed of 47.12 rad/sec. This range was inputted as a pseudorandom binary signal (PRBS) with a frequency range of 1-45Hz. The velocity of the cylinder was extracted from this position measurement to perform the system identification analysis. It was determined from a previous analysis, where only one pump was moving, that the system is fourth-order when using a velocity input and output. The output error (OE) parametric modeling method confirmed a fourth-order model for each operating mode. The system models for the normal mode, friction mode, and leakage mode are defined respectively as follows:

\[
y(s) = \frac{27.03s^3 + 9577s^2 + 1.87E6s + 1.65E8}{s^4 + 535.20s^3 + 7.38E4s^2 + 6.08E6s + 1.813E8}
\]  
\[
y(s) = \frac{11.29s^3 + 1883s^2 + 1.77E5s + 6.61E6}{s^4 + 151.80s^3 + 1.28E4s^2 + 6.14E5s + 1.07E7}
\]  
\[
y(s) = \frac{2.34s^3 + 910.60s^2 + 1.63E5s + 1.32E7}{s^4 + 116s^3 + 2.70E4s^2 + 1.41E6s + 3.521E7}
\]  

IV. EXPERIMENTAL RESULTS

To obtain the experimental results, note that the SVSF boundary layer was set to \(\psi = [100 1000 10000]^T\), and the SVSF ‘memory’ was defined as 0.1.

A variety of case studies were performed on the EHA setup, where the system ran under normal operating conditions and faults were injected throughout. The cases generally led to similar results. Consider the case where the system operates normally for the first five seconds, and then a leakage fault is introduced. The measurements were obtained and fed into the SVSF-IMM algorithm, as defined earlier by (7) through (23). For comparison purposes, the standard IMM (which uses the KF) was also used to generate results. The normal mode probability was calculated for this case, and is shown in the following figure.

![Fig. 5. The probability for the normal mode has been calculated, and is shown in the above figure. Note that 1 refers to 100%， and 0 refers to 0%.](image)

The input to the system started around 0.5 seconds, so the results prior to that should be ignored. Both the KF-IMM and the SVSF-IMM were able to obtain a very high probability of normal operation during the first five seconds. Once the fault was introduced, it took only a few samples for both methods to determine that a fault was occurring. The KF-IMM yielded roughly an 85% probability that the system was operating normally. The SVSF-IMM determined normal operation to be roughly 40%. Both methods dropped from a very high probability, thus indicating a fault.
Figure 6 shows the probability calculation of the system performing with a leak present. Both methods were able to correctly identify the presence of a fault at five seconds. The KF-IMM calculated roughly a 10% probability that a leak was occurring. Once steady-state was reached, the SVSF-IMM strategy determined the leakage fault with about a 60% probability. In terms of yielding a better estimate, the KF-IMM method actually yielded better results. However, when concerned with fault detection and diagnosis, the SVSF-IMM determined the presence of faults with a higher degree of probability. The difference is most likely attributed to the gain calculation of the SVSF, as shown by (4). The switching brings an inherent amount of stability, and is able to track model changes easily. In fact, when one decreases the boundary layer thickness, the magnitude of chattering is increased dramatically. Although this action reduces the quality of the estimate, it actually improves the ability of the SVSF-IMM to determine the presence of faults. For example, consider the same case but with the boundary layers reduced by 100 times.

The result of decreasing the boundary layer width is shown in the previous two figures. Essentially, chattering has been introduced into the estimation process, as demonstrated by the overshoot-type response shown between 5.5 and 6.5 seconds. The normal mode probability, at the inception of the fault, is shown to be roughly 20% for the SVSF-IMM, down from 40%. Likewise, the leakage mode probability increased to 80%, up from 60%. Therefore, the SVSF-IMM appears to yield a higher degree of probability when chattering is introduced into the estimation process. Although this comes at a cost of decreased performance, when one is concerned with fault detection and diagnosis, this is acceptable.

It is imperative to be able to detect and diagnose faults for the safe and reliable control of mechanical and electrical systems. The results of this paper indicate that the SVSF-IMM may be better suited for fault detection than the standard KF-IMM. Future work will include a sensitivity analysis for determining how close the different models can be before no reasonable distinction can be made by the SVSF-IMM.

V. CONCLUSIONS

This paper introduced a new interacting multiple model strategy by combining the standard IMM with the SVSF. The new strategy was applied on an experimental setup which included three types of operation: normal, friction, and leakage. System identification was performed for each operation in order to determine a corresponding model. Measurements taken from experimental trials were used by the KF-IMM and the SVSF-IMM to determine the operating mode probabilities. It was determined that the SVSF-IMM yielded better results for the purposes of fault detection and diagnosis. Furthermore, if chattering is introduced in the estimation process, a higher degree of mode probability may be determined.
APPENDIX

The following is a table of important nomenclature used throughout this paper:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j$</td>
<td>Number of filters (i.e., models used by IMM)</td>
</tr>
<tr>
<td>$x$</td>
<td>State vector or values</td>
</tr>
<tr>
<td>$z$</td>
<td>Measurement (system output) vector or values</td>
</tr>
<tr>
<td>$u$</td>
<td>Input to the system</td>
</tr>
<tr>
<td>$w$</td>
<td>System noise vector</td>
</tr>
<tr>
<td>$v$</td>
<td>Measurement noise vector</td>
</tr>
<tr>
<td>$A$</td>
<td>Linear system transition matrix</td>
</tr>
<tr>
<td>$B$</td>
<td>Input gain matrix</td>
</tr>
<tr>
<td>$H$</td>
<td>Linear measurement (output) matrix</td>
</tr>
<tr>
<td>$K$</td>
<td>Filter gain matrix (i.e., $KF$ or SVSF)</td>
</tr>
<tr>
<td>$P$</td>
<td>State error covariance matrix</td>
</tr>
<tr>
<td>$Q$</td>
<td>System noise covariance matrix</td>
</tr>
<tr>
<td>$R$</td>
<td>Measurement noise covariance matrix</td>
</tr>
<tr>
<td>$S$</td>
<td>Innovation (measurement error) covariance</td>
</tr>
<tr>
<td>$e$</td>
<td>Measurement (output) error vector</td>
</tr>
<tr>
<td>$\text{sat}(a)$</td>
<td>Defines a saturation of the term $a$</td>
</tr>
<tr>
<td>$p$</td>
<td>Probability transition matrix</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Likelihood term</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>SVSF ‘convergence’ or memory parameter</td>
</tr>
<tr>
<td>$\psi$</td>
<td>SVSF boundary layer width</td>
</tr>
<tr>
<td>$</td>
<td>a</td>
</tr>
<tr>
<td>$T$</td>
<td>Transpose of some vector or matrix</td>
</tr>
<tr>
<td>$\hat{x}$</td>
<td>Estimated vector or values</td>
</tr>
<tr>
<td>$k+1</td>
<td>k$</td>
</tr>
<tr>
<td>$k+1</td>
<td>k+1$</td>
</tr>
</tbody>
</table>

The following figure provides a more detailed circuit diagram of the EHA.

---

REFERENCES


---

*Fig. 9. The above figure shows more detail in regards to the EHA circuit diagram. Due to formatting constraints, the image may be difficult to view on paper.*