Spring 5-1-2017

A Novel Time-Domain Method of Fault Diagnosis in Induction Motors

Bryan J. Davis
bryan.j.davis@uconn.edu

A Novel Time-Domain Method of Fault Diagnosis in Induction Motors

Bryan Davis
Department of Electrical and Computer Engineering
University of Connecticut, Storrs, Connecticut, USA
bryan.j.davis@uconn.edu

***Abstract* – This paper introduces a novel time-domain method for detecting the four major types of induction motor faults without requiring complex signal processing or additional sensors other than those already in place for closed-loop motor control. This method artificially modulates the motor current feedback signal with a perturbed frequency that oscillates about the characteristic frequency of the fault in question. After filtering out the unwanted frequency components that result from the modulation, the selected fault-indicative component is a very slow sinusoid that is only present when the fault is present. The method looks for this fault-indicative component by monitoring the time since the last zero crossing, while also adapting the modulating signal in real-time based on the motor speed and torque. The perturbations of the modulation frequency are employed to accentuate the differences between fault and no-fault conditions and increase detection speed. Simulations are conducted in Simulink to validate the accuracy and speed of this detection method for each fault type under a variety of operating conditions and commanded speeds. The proposed method is capable of correctly detecting all the fault types, while offering exceptional detection speed and the ability to detect multiple faults concurrently.**

I. INTRODUCTION

A. Induction Motors and Faults

Induction motors are the workhorses of the industrial world, serving as the integral components in the conversion of electrical to mechanical energy in most manufacturing processes. In addition to industry, induction motors' use extends to heating and ventilation, sewage and irrigation, and elevators. Regardless of their application, maintaining a healthy motor is crucial for proper and safe operation, and, consequently, damages to the motor should be recognized and remedied quickly. If a fault in an induction motor is not detected in its early stages, it can spread and become more serious, not only requiring a longer and more costly downtime for repair, but also making a dangerous catastrophic failure more likely. [1]

Faults within induction motors can be categorized into four major types: broken rotor bar fault (BRBF), air-gap eccentricity fault (EF), bearing fault (BF), and stator short winding fault (SSWF). The windings of the rotor are typically in the form of bars rather than wires, so a BRBF is, as the name suggests, a split or hole in a rotor bar. Often caused by thermal stress due to high induced current as well as high ambient operating temperatures, a single BRBF will likely spread to other rotor bars and become more serious, if not mitigated rapidly [2].

An EF occurs when the air gap between the rotor cage and the stator windings is not constant. A misaligned shaft can cause the air gap to be constant for a given angular position within

the motor, but different when compared to the spacing at other angular positions, resulting in static eccentricity. Alternatively, a bent rotor shaft or oval core can cause the air gap to vary periodically with time at a given angular position, resulting in dynamic eccentricity (DE); or static and dynamic eccentricity can combine to form mixed eccentricity (ME). A further implication of eccentricity is the addition of frequency components, called principal slot harmonics (PSH), to the stator current feedback signal. These arise from the rotor bars themselves, which form slots in the rotor cage and cause variance in the air gap [3]. Generally speaking, because rotor slots are present even in healthy motors, PSH always exist, but the injection of an EF adds PSH of other orders to the current feedback [3]. [1]

The bearings in an induction motor support the rotor shaft within the housing and minimize friction during rotation of the shaft. The rolling elements in the bearings, either balls or cylinders, sit between the inner race, which is secured to the shaft, and the outer race, which is secured to the housing. A BF involves damage, such as cracks, to any of these components and can result simply from fatigue and wear over time but can be accelerated by factors, such as lack of lubrication, high load, and long run times. [2]

Finally, a SSWF results from an electrical short within the windings of the stator. A turn-to-turn SSWF, the least severe of the three SSWFs, occurs when the insulation in a single phase winding is worn, causing adjacent turns in the coil to short. A phase-to-phase SSWF happens when two separate phase windings are shorted at some point within the windings, and a phase-to-ground SSWF happens when a point in the winding of one phase is shorted to ground, usually the housing of the motor. A variety of thermal, environmental, and mechanical stresses can cause a SSWF [1]. [4]

Myriad fault detection and diagnosis (FDD) methods are utilized in practice, and even more are proposed in the literature, but the foundation behind all the methods is the same: identify and detect a characteristic change in a feedback signal that occurs only if a fault appears. The signal upon which the fault analysis is done is usually the stator current or the vibration of the motor. Generally, each fault produces an abnormal feature in the current or vibration signal that is characteristic only to that fault; if this feature can be detected, then the fault can be diagnosed when present. The four major methods of fault diagnosis are time-domain, frequency domain, both time and frequency domain, and artificial intelligence. [2]

The benefits of the method proposed in this paper over previously used methods are fourfold. First, the detection of the fault is simple and requires minimal signal processing. Second, no additional sensors, besides those already in place, are needed. Third, all major types of faults can be detected, even if they are present concurrently. Lastly, this method offers very fast detection, minimizing the time that the machines run while damaged. Previous methods may offer some of these benefits, but none offer all of these beneficial effects at the same time. Due to the widespread use of induction motors, this quick, comprehensive, and inexpensive fault detection method is paramount for safety as well as industrial efficiency.

B. Unsuccessful Time-Domain Attempts

As previously mentioned, each fault adds an abnormality to the stator current signal that is characteristic only to that fault. In particular, in the frequency domain, the four major faults add

frequency components to the stator current as predicted by the equations in Table I, where k_1 to k_7 are integers that are related to the harmonic order of fault signals, f_e is the synchronous frequency; s and P are machine slip and number of poles, respectively, N_b and N_s are the number of bearing rolling elements and rotor slots, respectively, and N_d and N_w are the orders of rotating eccentricity and stator magneto-motive force harmonics, respectively [1]. If one fault is present in the current feedback, then the stator current $y_c(t)$ would be

$$y_c(t) = A\sin(\omega_e t + \theta_{e0}) + a\sin((\omega_{fm} + 2\pi d)t + \theta_{f0}), \tag{1}$$

where A is the fundamental component magnitude, a is the fault component magnitude ($A \gg a$), ω_e is the fundamental frequency (in rad/sec), ω_{fm} ($=2\pi f_{fm}$) is the fault characteristic frequency as determined by Table I, and θ_{e0} and θ_{f0} are the initial phases of each sinusoid. In practice, the real fault signal frequency varies slightly from the calculated characteristic frequency by a value d (typically ±1 Hz), which arises from factors, such as machine vibrations and instrument inaccuracies [1].

Table I. Characteristic fault frequencies for the four major fault types [1]

Fault	Fault Characteristic Frequency Components	
EF	Dynamic and mixed EF	$f_{ef1} = f_e\left[1 \pm \frac{2k_1(1-s)}{P}\right]$
	Principal slot harmonics	$f_{ef2} = f_e\left[\frac{2(k_2 N_s \pm N_d)(1-s)}{P} \pm N_w\right]$
BF	Simplified inner-race/ outer-race BF	$\begin{bmatrix} f_{ir} \\ f_{or} \end{bmatrix} = f_e + \begin{bmatrix} 0.6k_3 N_b f_{rm} \\ 0.4k_4 N_b f_{rm} \end{bmatrix}$
BRBF	$f_{brbf} = f_e[1 \pm 2k_5 s]$	
SSWF	$f_{sswf} = f_e\left[\frac{2k_6(1-s)}{P} \pm k_7\right]$	

If the current feedback is modulated with a sinusoid of angular frequency ω_m that matches the actual frequency of the present fault, then the fault can be easily isolated because a DC component will result in the product. Let the modulating signal $y_m(t)$ with amplitude X be given as

$$y_m(t) = X\sin(\omega_m t). \tag{2}$$

Then, using trigonometric identities, the product $y(t)$ of the modulation will be

$$\begin{aligned} y(t) &= y_c(t) \times y_m(t) \\ &= \frac{AX}{2}\cos\left[(\omega_e - \omega_m)t + (\theta_{e0} - \theta_{m0})\right] \\ &\quad - \frac{AX}{2}\cos\left[(\omega_e + \omega_m)t + (\theta_{e0} + \theta_{m0})\right], \\ &\quad + \frac{aX}{2}\cos(\theta_{f0}) \\ &\quad - \frac{aX}{2}\cos\left[2(\omega_{fm} + 2\pi d)t + (\theta_{f0})\right] \end{aligned} \tag{3}$$

where the DC component has a value of $\frac{aX}{2}\cos\left(\theta_{f0}\right)$ and the other three components have relatively high frequencies. This DC component is only present when a fault is present, so isolation of DC means that a fault can be flagged. The goal of my early attempts were thus to find and detect this DC component by modulating at frequencies close to the characteristic frequency of the fault in question.

The first method I attempted involved continually recalculating the modulating frequency until a DC, or almost DC, component arose. If ω_m did not match the actual fault frequency (in rad/sec) exactly, then the third component in (3) would be $\frac{aX}{2}\cos\left[\left(\left(\omega_{fm}+2\pi d\right)-\omega_m\right)t+\left(\theta_{f0}\right)\right]$ instead of DC. Using the approximate time between zero crossings of this sinusoid, ω_m could be adjusted repeatedly until $\omega_m=\omega_{fm}+2\pi d$ and the DC is isolated. The largest issue with this method, however, is the need to wait for zero crossings that successively become farther apart as ω_m is adjusted, thus resulting in a very long detection time. Moreover, ripples from the other components in (3) would greatly affect the time between zero crossings and thus the convergence of ω_m to the actual fault frequency. Additionally, the *d* value is not necessarily constant with time, so the DC component would disappear as soon as *d* changes.

Next, I attempted to create and isolate the DC component by modulating the current feedback signal with a sweep of frequencies. When the modulating frequency matched the actual fault frequency exactly, the DC component would arise and the frequency of the present fault could be determined to then ascertain the type of fault that is present. However, as the modulating frequency sweeps past the fault frequency, the DC component only occurs for exactly one moment in time: the time of exact frequency matching. Otherwise, the third component in (3) is sinusoidal. Due to the presence of the three other components in *y(t)* and only a single moment of DC, this DC component is very difficult to recognize and detect in the *y(t)* signal, even if the other three components are largely attenuated. Even though this specific method was unsuccessful, the concept of modulating with a varying, rather than constant, frequency was the foundation behind the development of perturbed-frequency adaptive modulation (PFAM).

These unsuccessful attempts were valuable because they first illustrated the importance of filtering out the three other components in (3) so that they interfere minimally with the fault-indicative component. Secondly, they demonstrated that this FDD method should not search for precisely DC but rather for a very slow sinusoid, since it is difficult to find and maintain the exact fault frequency.

II. Overview of Adaptive Modulation

Now, let the current feedback be modulated with a sinusoid that has the characteristic angular frequency ω_{fm} of the fault in question; the modulating signal is then

$$y_m(t)=X\sin\left(\omega_{fm}t\right). \tag{4}$$

As in (3), the product *y(t)* of the modulation will be

$$
\begin{aligned}
y(t) &= y_c(t) \times y_m(t) \\
&= \frac{AX}{2}\cos\left[\left(\omega_e - \omega_{fm}\right)t + \left(\theta_{e0} - \theta_{m0}\right)\right] \\
&\quad - \frac{AX}{2}\cos\left[\left(\omega_e + \omega_{fm}\right)t + \left(\theta_{e0} + \theta_{m0}\right)\right], \\
&\quad + \frac{aX}{2}\cos\left[\left(2\pi d\right)t + \left(\theta_{f0}\right)\right] \\
&\quad - \frac{aX}{2}\cos\left[\left(2\omega_{fm} + 2\pi d\right)t + \left(\theta_{f0}\right)\right]
\end{aligned}
\tag{5}
$$

which contains four clear frequency components. As before, the third component in (5), with frequency *d*, is chosen as the fault-indicative component because it is a very slow sinusoid with a frequency far away from the three other components, and it is present only when a fault is present. When a fault is not present in $y_c(t)$, the third and fourth components in (5) do not exist. This modulation scheme, recently developed at the University of Connecticut Advanced Power Electronics and Electric Drives Laboratory (APEDL) and described in [1], is called adaptive modulation because as the machine conditions (such as slip, speed, and torque) change, the modulating frequency adapts to match the changing characteristic fault frequencies based on the equations in Table I.

If various signal processing techniques are applied to $y(t)$ to filter out all but the third frequency component, then the presence of the third frequency component in the post-processed modulated signal $y_{post}(t)$ will indicate that a fault has occurred. The basis of the adaptive modulation method is the use of the time between zero crossings of this $y_{post}(t)$ signal to determine if the fault-indicative component is present. If no fault is present, $y_{post}(t)$ will consist of noise centered around zero, for which the time between zero crossings will be very small. However, if a fault is present, the fault-indicative component, with frequency *d*, will arise in $y_{post}(t)$ and the time between zero crossings will increase to about $\frac{1}{2d}$, as shown in Fig. 1. Now, let a threshold T_{t_th} be defined as

$$
T_{t\_th} = \frac{1}{2d_{\max}}, \tag{6}
$$

where d_{max} is the most the actual fault frequency can drift from ω_{fm} while still being considered a fault. In $y_c(t)$, $d = d_{max}$ would result in the highest frequency of the fault-indicative component and thus the shortest time between zero crossings in $y_{post}(t)$ when a fault is present. Therefore, as long as the time between zero crossings is greater than or equal to T_{t_th}, the fault-indicative component must be present and a fault can be flagged. Otherwise, no fault is triggered. [1]

For a low *d*, the time between zero crossings could be very large, so this method does not wait for the next zero crossing because this could lead to very long detection times. Instead, this method defines T_t, which is the difference between the current simulation or experimental runtime and the time of the most recent zero crossing. Then, as soon as T_t surpasses T_{t_th}, the fault flag *r* goes to 1 to indicate a fault (otherwise $r = 0$ to indicate no fault). Although this

technique increases the average speed of detection t_d after a fault occurs, t_d is still at least T_{t_th} and is therefore no less than 0.5 seconds for d_{max} = 1 Hz. [1]

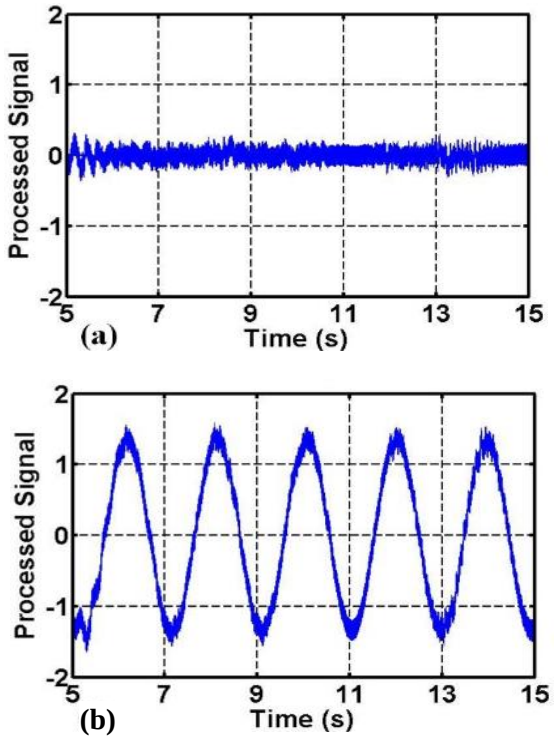

Fig. 1. $y_{post}(t)$ signal without (a) and with (b) fault present [1]

III. Perturbed-Frequency Adaptive Modulation (PFAM)

A. Motivation

In the described method of general adaptive modulation, the basis behind the detection of a fault is the confirmation that the time between zero crossings is greater than T_{t_th}. As a result, the model must wait at least time T_{t_th} before declaring a fault. However, to obtain an even faster detection time, T_{t_th} can be lowered to the maximum time between zero crossings (defined as T_{zm}) of the non-fault $y_{post}(t)$ signal. Based on this definition of T_{zm}, if the time between zero crossings surpasses T_{zm}, then a fault must be present because the time between zero crossings is limited to T_{zm} while there is no fault. Therefore, if $T_{t\_th} = T_{zm}$, then, as before, the instant T_t surpasses T_{t_th}, the model will flag a fault, although now this reaction time will be significantly lower because

$$T_{zm} < \frac{1}{2d_{\max}}, \tag{7}$$

as was used to define T_{t_th} previously. The no-fault $y_{post}(t)$ signal for dynamic/mixed EF (DE/ME) and the maximum time between zero crossings used to define the fault's T_{zm} are shown in Fig. 2. T_{zm} in this instance is less than T_{t_th}'s prior value of 0.5s, but there are still considerable gaps between zero crossings, thus limiting the detection time. Therefore, this limitation introduces the need for PFAM.

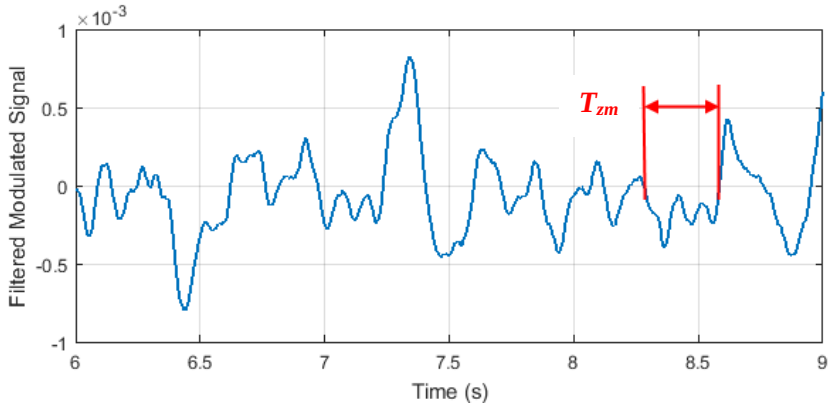

Fig. 2. The $y_{post}(t)$ signal for DE/ME when no fault is present. T_{zm} = 302 ms

B. PFAM

Let the modulating signal now be defined

$$y_{pm}(t) = X\sin\left(2\pi\left[f_{fm} + \Delta f\sin(2\pi f_p t)\right]t\right), \tag{8}$$

where f_{fm} is the center modulating frequency as well as the ideal fault characteristic frequency from Table I, Δf *is* the variation in modulating frequency, and f_p is the frequency of the perturbation. Before, the modulating frequency was a constant, although it updated based on the real-time speed, slip, and torque of the motor. Now, the modulating frequency varies from $f_{fm} + \Delta f$ to $f_{fm} - \Delta f$, as shown in Fig. 3, while also updating with changes in motor speed, slip, and torque. This is similar to the concept of swept sine waves used in system identification. The adaptive modulation on the current feedback now results in

$$\begin{aligned} y(t) &= y_c(t)\times y_{pm}(t) \\ &= \frac{AX}{2}\cos\left[\left(\omega_e - \omega_{fm} - 2\pi\Delta f\sin(2\pi f_p t)\right)t + \theta_{e0}\right] \\ &\quad - \frac{AX}{2}\cos\left[\left(\omega_e + \omega_{fm} + 2\pi\Delta f\sin(2\pi f_p t)\right)t + \theta_{e0}\right] \\ &\quad + \frac{aX}{2}\cos\left[2\pi\left(d - \Delta f\sin(2\pi f_p t)\right)t + \theta_{f0}\right] \\ &\quad - \frac{aX}{2}\cos\left[\left(2\omega_{fm} + 2\pi d + 2\pi\Delta f\sin(2\pi f_p t)\right)t + \theta_{f0}\right] \end{aligned} \tag{9}$$

As before, the third component, with frequency $\left(d - \Delta f\sin(2\pi f_p t)\right)$, is selected as the fault-indicative component, and the other three undesired components are filtered out as much as possible. The signal processing stage of PFAM achieves this with three filters. The $\left(2\omega_{fm} + 2\pi d + 2\pi\Delta f\sin(2\pi f_p t)\right)$ component is attenuated with a LPF; the $\left(\omega_e - \omega_{fm} - 2\pi\Delta f\sin(2\pi f_p t)\right)$ component is attenuated with the same LPF as well as a tunable bandstop filter set to block the $\left(\omega_e - \omega_{fm}\right)$ angular frequency; and the $\left(\omega_e + \omega_{fm} + 2\pi\Delta f\sin(2\pi f_p t)\right)$ component is attenuated with the same LPF and a different tunable bandstop filter set to block the $\left(\omega_e + \omega_{fm}\right)$ angular frequency.

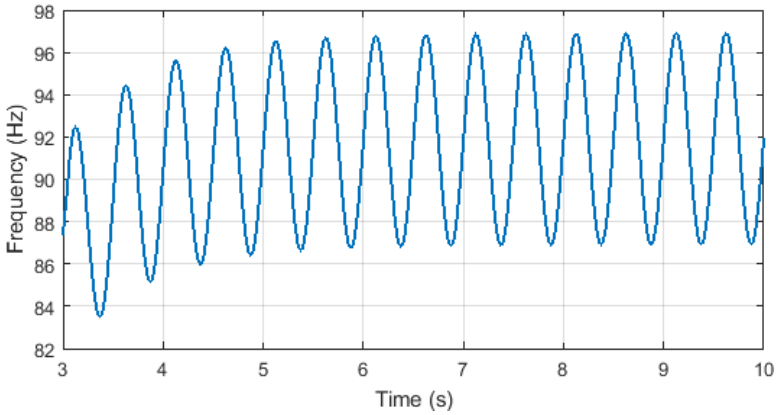

Fig. 3. Modulating Frequency for DE/ME with Δf = 5 Hz and f_p = 2 Hz

When there is no fault present, adaptive modulation on the current feedback results in

$$\begin{aligned} y(t) &= y_c(t) \times y_{pm}(t) \\ &= \frac{AX}{2}\cos\left[\left(\omega_e - \omega_{fm} - 2\pi\Delta f \sin(2\pi f_p t)\right)t + \theta_{e0}\right] \\ &\quad - \frac{AX}{2}\cos\left[\left(\omega_e + \omega_{fm} + 2\pi\Delta f \sin(2\pi f_p t)\right)t + \theta_{e0}\right] \end{aligned}, \tag{10}$$

where, in this case, both components are *y(t)* are greatly attenuated by the three filters. It is important to point out, however, that because the bandstop filters are tuned exactly to $\left(\omega_e \pm \omega_{fm}\right)$, the two components in (10) will be decreasingly attenuated as the modulating frequency approaches $\left(f_{fm} \pm \Delta f\right)$. The more the modulating frequency varies from the center frequency of f_{fm}, the more these relatively high frequency components show up in the post-processed *y(t)*. The consequence of this perturbed frequency is thus an increased number of oscillations in the processed modulated signal. Because these oscillations interfere with the fault-indicative signal, the goal of PFAM is to limit these higher frequency components, but only to a certain degree, since they also have the benefit of decreasing T_{zm} and thereby decreasing detection time, as will be shown below.

C. Fault Detection

Figure 4 demonstrates how applying a perturbed frequency, with Δf = 5 Hz and f_p = 2 Hz, to the same DE/ME under the same machine conditions as Fig. 2 causes T_{zm} to decrease from 304 ms to 24 ms. Theoretically, a fault could then be detected in time

$$t_d = T_{zm} + t_r, \tag{11}$$

where t_d is how long the proposed method takes to flag the fault after it is injected, and t_r is the fault reaction time, or the time to the last zero crossing after fault-injection before the fault takes effect, as depicted in Fig. 5. Here, t_r can vary depending on the time of fault-injection and the relative phase of *y(t)*. In Fig. 5, after a DE/ME is injected at 7.3615s, the $\left(d - \Delta f \sin(2\pi f_p t)\right)$ component arises and becomes dominant. Although the frequency perturbations cause smaller

additional oscillations in this fault-indicative signal, the distance between zero crossings is still much larger than T_{zm} and a fault is thus flagged. The parameters Δf and f_p can be tuned to adjust the degree of additional oscillations present. Although f_p has minimal effect, a greater Δf results in more oscillations, as the frequencies further from $\left(\omega_e \pm \omega_{fm}\right)$ are attenuated less by the bandstop filters. Δf should be set as high as possible to increase the number of oscillations and limit T_{zm} while a fault is not present, while remaining low enough that the fault-indicative component is still distinguishable.

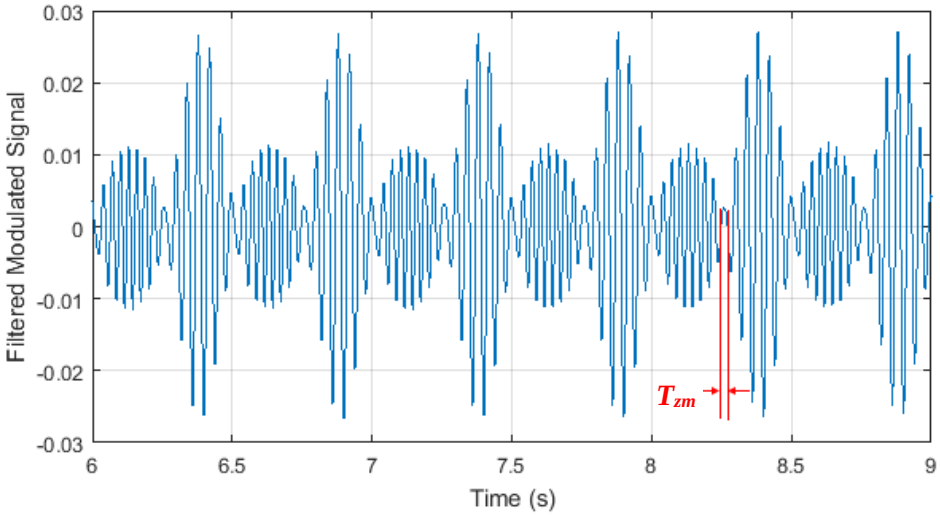

Fig. 4. The no-fault $y_{post}(t)$ signal for DE/ME with PFAM applied (Δf = 5 Hz and f_p = 2 Hz). T_{zm} = 24 ms

Although the addition of frequency perturbations allows for much quicker detection times, it necessitates additional stages to the fault detection system. First, due to the added oscillations in the fault-indicative component, an additional fault condition is included to ensure that, when a fault is present, a limited number of quick oscillations do not remove the fault flag. Let this additional fault condition be called Fault Condition 2 (FC2), and let the previously described fault condition, which checks if $T_t > T_{zm}$, be called FC1. If the fault has already been flagged and these quick oscillations occur in the fault-present $y_{post}(t)$ as the larger overall $y_{post}(t)$ is crossing zero (as shown in Fig. 6), several smaller oscillations with zero crossing times less than T_{zm} would otherwise indicate that the fault has disappeared by FC1. Here, FC1 would drop from logic 1 (indicating a fault) to logic 0 (indicating no fault) during these rapid zero crossings. However, FC2 will not allow the fault flag to disappear after a fault has already been flagged unless more than n_z zero crossings occur with less than time T_{zm} passing between each subsequent zero crossing. This n_z is currently a user-defined value but can potentially be calculated by the model based on the machine parameters and Δf and f_p values. If n is the number of zero crossings after FC1 drops back to 0, FC2 will only indicate that the fault has disappeared if $n \geq n_z$ and FC1 = 0. If FC1 = 1, FC2 will also always equal 1, as no unwanted quick oscillations would be interfering with the fault-indicative signal in this case.

Secondly, the determination of T_{zm} is difficult due to the large variability of machine conditions and signal parameters that would influence the shape of the no-fault processed $y(t)$. Given the no-fault $y_{post}(t)$, as in Fig. 7, one could measure the longest time between zero

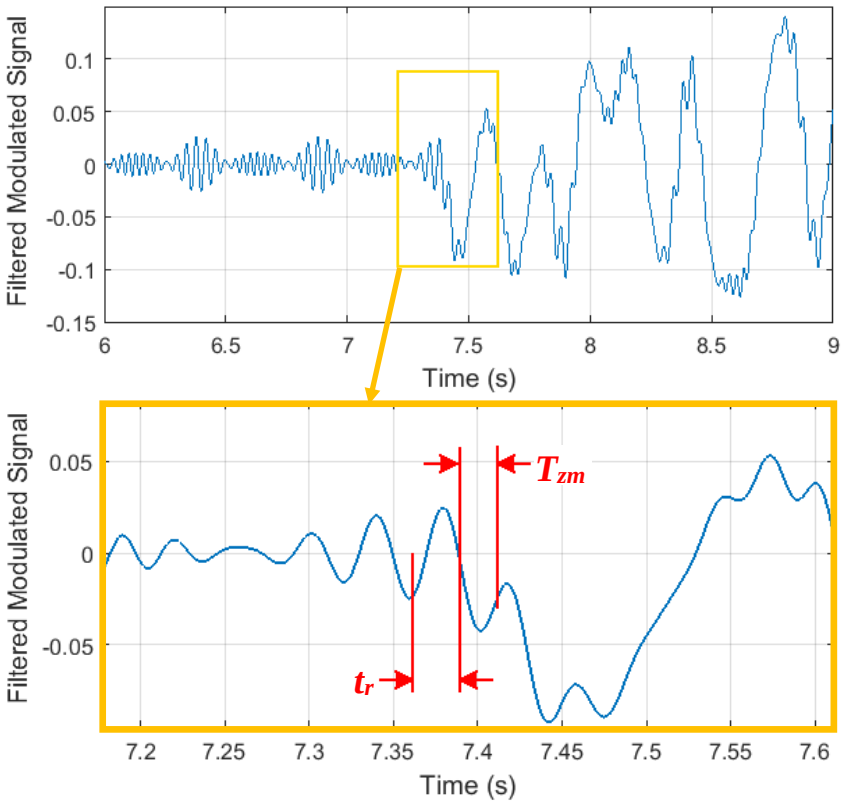

Fig. 5. The $y_{post}(t)$ signal with DE/ME injected at 7.3615s. t_r = 27 ms and T_{zm} = 24 ms, giving t_d = 51 ms

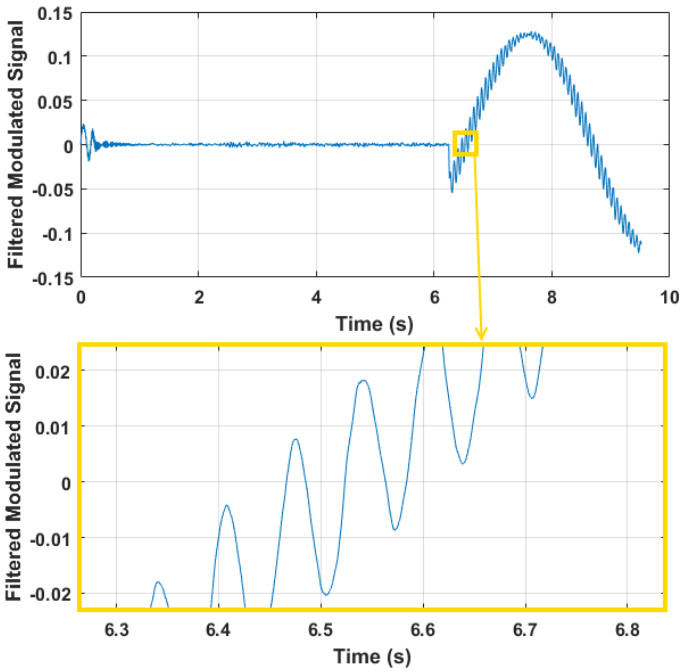

Fig. 6. Exemplary quick oscillations about zero after a PSH fault is injected at 6.2852s

crossings, but this is not necessarily the true value of T_{zm} because a larger zero crossing gap could theoretically occur outside the scope of the given plot. T_{zm} should be set as low as possible to increase detection speed, but if it is set too small, then the proposed method becomes more

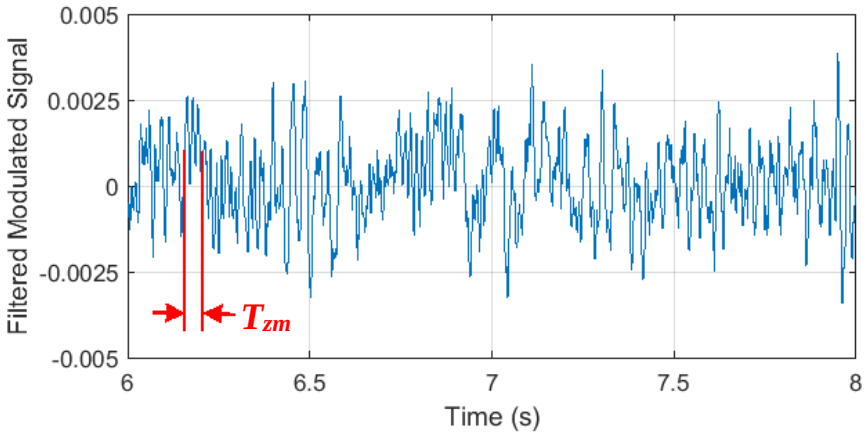

Fig. 7. The $y_{post}(t)$ signal with a PSH fault present (Δf = 5 Hz and f_p = 2 Hz). Here, T_{zm} = 49 ms.

prone to false flags. On the other hand, if T_{zm} is set too large, then, even when a fault is present, the time between zero crossings could remain less than T_{zm} and a fault could be ignored. Due to the difficulty in finding the proper balance for T_{zm}, a third fault condition, FC3, is added to add reliability to the proposed method even if T_{zm} is set too low. With FC3, if no fault is flagged and more than T_{zm} time units pass since the last zero crossing, the fault flag will not appear immediately (although FC1 will go to 1). Instead, the proposed method will wait an additional user-defined time T_w after FC1 flags a fault; after time T_w, FC3 will check the status of FC1 again and only flag a fault if FC1 is *still* indicating a fault condition. This greatly decreases the chances of a false flag, although the detection time is then increased by T_w.

D. Consider DC Offset, Noise and Harmonics in the Stator Current Feedback

It is common for the current feedback to contain noise, harmonics, and a DC offset. With PFAM, a DC component with magnitude B adds a $BX\sin\left(\left[\omega_{fm}+2\pi\Delta f\sin(2\pi f_p t)\right]t\right)$ component to the modulated signal. However, this component is greatly attenuated by the LPF and has negligible effect on $y_{post}(t)$, as seen in Fig. 7, where a DC component with B = 1 has been included in the current feedback. Moreover, the noise and harmonic contents in the current feedback add AC components to the modulated signal, but these components' frequencies are significantly higher than the desired $\left(\omega_e-\omega_{fm}\right)$ component and are thus expected to be greatly attenuated by the LPF. Furthermore, the effect of any remaining high-frequency noise or harmonics in the processed $y(t)$ should be limited by FC2, which compensates for up to n_z small ripples as the larger $y_{post}(t)$ crosses zero. Further testing will need to be completed to confirm the robustness of this method to noise and higher harmonics.

E. Advantages and Disadvantages

As is the case with general adaptive modulation, PFAM offers a simple FDD method that requires minimal signal processing and no additional sensors, besides those already in place for control purposes [1]. More importantly, PFAM can detect faults quickly, even faster than general adaptive modulation. As will be shown in the Simulation Verification section, this method can detect a fault as quickly as 78.6 ms, thus allowing faults to be detected and remedied in their

incipient stages. Moreover, the Simulation Verification section will illustrate how this method can correctly detect multiple faults that are present concurrently, and how, due to the adaptive nature of this modulation, this method can correctly detect faults for a variety of commanded speeds without any adjustment of the model parameters.

This increase in fault detection speed does not come without limitations, however. First, T_{zm}, T_w, and n_z must be set for each fault type (although n_z can be constant for all faults or can be calculated based on an algorithm of model parameters, as previously mentioned). Nonetheless, companies that rely on induction motors for production or safety-critical applications will likely find the extra off-line calibration time to be worth the diminished on-line fault detection time. Second, the determination of the optimal values of Δf and f_p is very difficult. For instance, if Δf is too low, T_{zm} remains too high due to lack of additional oscillations and detection is slow. On the other hand, if Δf is too high, too much ripple exists in the fault-indicative signal and the fault cannot be detected. After many hours of adjusting, Δf = 2 Hz and f_p = 15 Hz proved to be effective settings that worked for every fault type, but they are not necessarily the optimal values leading to the lowest T_{zm}, while also reducing ripples in a faulty state. However, the APEDL group hopes to conduct further research on PFAM by utilizing extremum seeking control (ESC) to force Δf and f_p to converge to their optimal values.

Additionally, as mentioned, a direct consequence of employing additional oscillations to decrease zero crossing time is the additional ripples in the fault-indicative signal, thus reducing fault sensitivity. As will be shown in the Simulation Verification section, fault magnitudes less than 3.08% of the fundamental magnitude may not be properly detected. This is because the extra ripple is so high relative to the fault-indicative signal's magnitude that the ripples appear to be centered around zero even when there is a fault, or the extra ripples cause FC2 to drop to 0 repeatedly. Lastly, PFAM is more prone to false flagging due to unwanted frequency components in the current feedback that are relatively close to the characteristic frequency of a fault. Such a frequency component could be modeled as an actual fault with $d > d_{max}$. Ideally, PFAM would ignore this frequency because it would have a drift higher than d_{max}, but, in fact, it only ignores the frequency if

$$d > \frac{1}{2T_{zm}}, \tag{12}$$

because then the time between zero crossings of the fault-indicative signal (every half period) would be less than T_{zm} for that fault. The remedy to this disadvantage would be to wait for the next zero crossing after the fault is injected and then approximate whether d is below d_{max}, but the method would consequently lose its chief advantage of speed.

IV. Simulation Verification

A. Model

The proposed method is verified in MATLAB/Simulink using a closed-loop indirect field oriented control (IFOC) model for an inverter-fed 1.5HP induction machine, as shown in Fig. 8 [1]. The PFAM model described in this paper is contained in the Fault Detection block and is

shown in detail in Fig. 9, using the detection of DE/ME as an example. In the full PFAM model, there would be 6 modulation branches for each of the 6 characteristic fault frequencies, each looking specifically for their own assigned fault. A PLL block is applied on the current feedback to obtain the synchronous frequency f_e. Then the theoretical fault frequency of DE/ME is calculated using f_e, the speed feedback, and the equation in Table I, where k_1 is selected to be 1. Prior to running the simulation, a MATLAB m-file is run to obtain various parameters for the

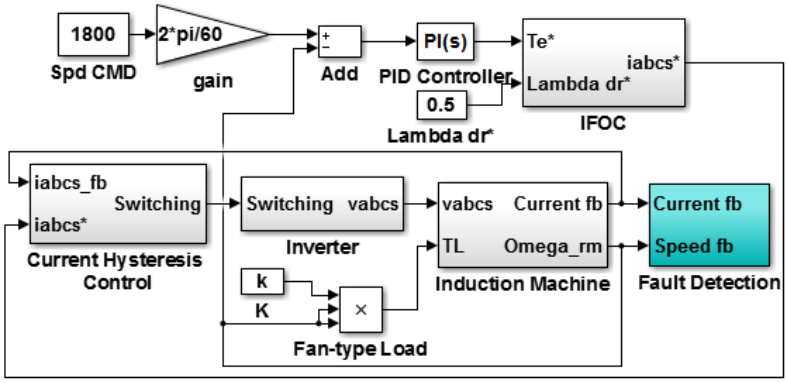

Fig. 8. High-level block diagram of the inverter-fed IFOC induction machine with added fault detection [1]

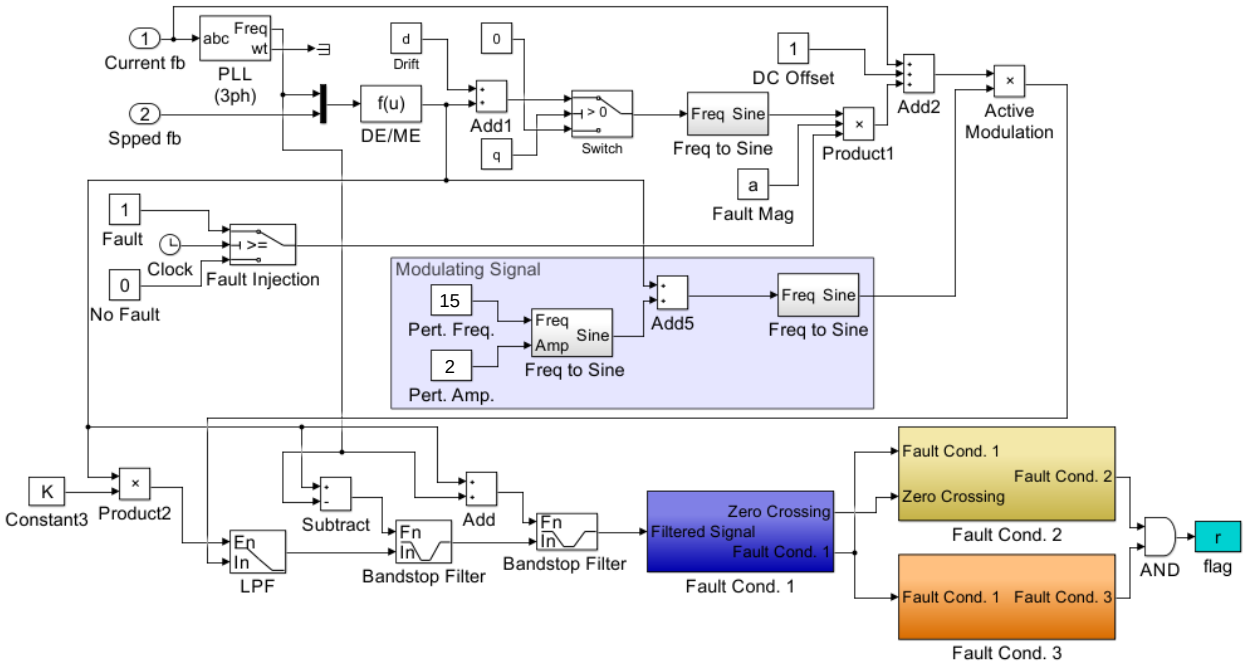

Fig. 9. The PFAM subsystem of the Fault Detection block in Fig. 8

model as well as the random drift d, the randomly-assigned fault trigger q, and random time t_1. d is assigned to be no more than ± 1Hz from the fault characteristic frequency, and the fault is only passed if $q=1$. The fault injection switch passes a 1 to Product1 to activate the fault after $(6+t_1)$ seconds have passed, where 6s is approximately the time necessary to reach steady-state after

startup. The output of the Add2 block is thus a virtual imitation of a real motor's current feedback, where DC could be present and a fault component could arise at any time.

Next, the modulating signal and Add2 output are multiplied together, the result of which passes through the LPF and two bandstop filters. This processed modulated signal is then checked by fault conditions 1, 2, and 3. Because FC2 and FC3 are dependent on FC1, only the outputs of FC2 and FC3 are passed through a logical AND to determine the fault flag value *r*. The correctness of the proposed method is determined by confirming that the *q* and *r* values ultimately match. Lastly, to ensure that the presence of an outer race BF (ORBF) does not interfere with the detection of a SSWF, or vice versa (the two are very close in characteristic frequency: 153.9Hz and 157.9Hz at 1800 RPM), an additional bandstop filter is added specifically to the detection branch of these two faults. This filter is tuned to $\left(\omega_{fm1} - 2\pi\left[f_{fm2} + \Delta f \sin(2\pi f_p t)\right]\right)$ to attempt to filter out the slow frequency component that arises as the difference between the potentially present nearby fault and the instantaneous frequency of the current modulating signal. The subscript 1 indicates the characteristic frequency of the first fault (SSWF or ORBF) and the subscript 2 indicates the characteristic frequency of the other fault.

B. Simulation Results

The proposed fault detection model is run for each individual fault type with Δf = 2 Hz, f_p = 15 Hz, a = 0.25, n_z = 10, and t_1 = 0.3424s, thus injecting the faults at 6.3424s. For an inner race BF (IRBF), the subsequent fault injection trigger and fault flag signals are shown in Fig. 10a, which demonstrates that the trigger signal goes to 1 at 6.3424s, but the flag does not jump to 1 until 6.5560s, giving a t_d of 213.6 ms. The detection time is directly affected by the chosen T_{zm} and T_w times as well as the reaction time t_r of the fault; all of these values for each fault type are summarized in Table II. The corresponding $y_{post}(t)$ signal is shown in Fig. 10b to illustrate the difference in the signal with and without a fault. By looking at the plot of this signal, it is seen that t_r in this case is 3.9 ms. Thus, the expected t_d is

$$\begin{aligned} t_d &= t_r + T_{zm} + T_w \\ &= \left(3.9 + 105 + 105\right)\ \text{ms} = 213.9\ \text{ms}, \end{aligned} \tag{13}$$

which matches the simulation t_d of 213.6 ms almost exactly (the slight error here, as well as in the other values in Table II, is due to the lack of precision in the Simulink scope window). While conducting these simulations for each fault, there were no false flags prior to the fault injection, and the flag remained at 1 during the remainder of the simulation after fault-injection, as expected.

To test the method's ability to detect multiple faults concurrently, the latter half of the subsystem shown in Fig. 9 is replicated to create a second modulation branch that modulates the same Add2 block output (but with a second fault frequency added). In this dual-branch model, the injected faults and f_{fm} of the modulating signals are adjusted to achieve every permutation of the 6 faults, and then the model is run with the parameters in Table II to see if both of the concurrently injected faults are detected in their respective modulation branch. After

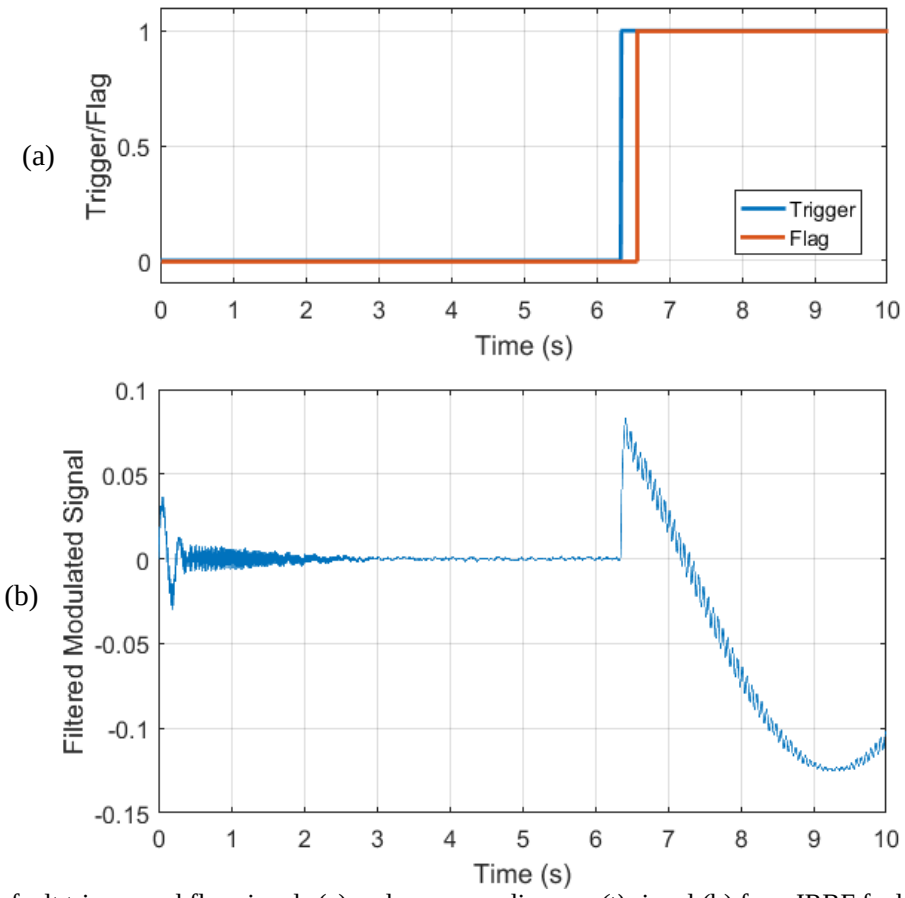

Fig. 10. The fault trigger and flag signals (a) and corresponding $y_{post}(t)$ signal (b) for a IRBF fault injected at 6.3424s. t_d = 213.6 ms

Table II. Fault detection time, and its components, for each fault for Δf = 2 Hz, f_p = 15 Hz

Fault Type	f_{fm} **(Hz)**	T_{zm} **(ms)**	T_w **(ms)**	t_r **(ms)**	t_d **(ms)**
DE/ME	91.95	40	40	24.7	104.6
PSH	931.9	39	39	1	78.6
BRBF	65.85	57	57	32.3	146.6
SSWF	153.9	195	10	7	211.6
IRBF	205.9	105	105	3.6	213.6
ORBF	157.9	195	10	1.2	206.6

completing this test, it was concluded that PFAM can indeed detect multiple faults concurrently. For every permutation, both faults were quickly flagged after injection and the flag remained for the duration of the simulation.

Furthermore, to confirm that PFAM is capable of producing the same results at low speed, the commanded speed shown in Fig. 8 is adjusted to 600 RPM. Then, without adjusting the parameters shown in Table II from the 1800 RPM trials, the simulation is run twice for each type of fault (a new random drift is generated for each fault on the second trial). At this low speed, all faults were correctly and quickly detected for both trials, with one exception. In the

second IRBF trial, the fault was not flagged because the low randomly-generated drift of $d = -0.1225$ Hz resulted in a very slow sinusoid with high ripple that forced FC2 to 0 every time the larger fault-indicative sinusoid crossed zero, similar to in Fig. 11. n_z would simply have to be increased to avoid this error.

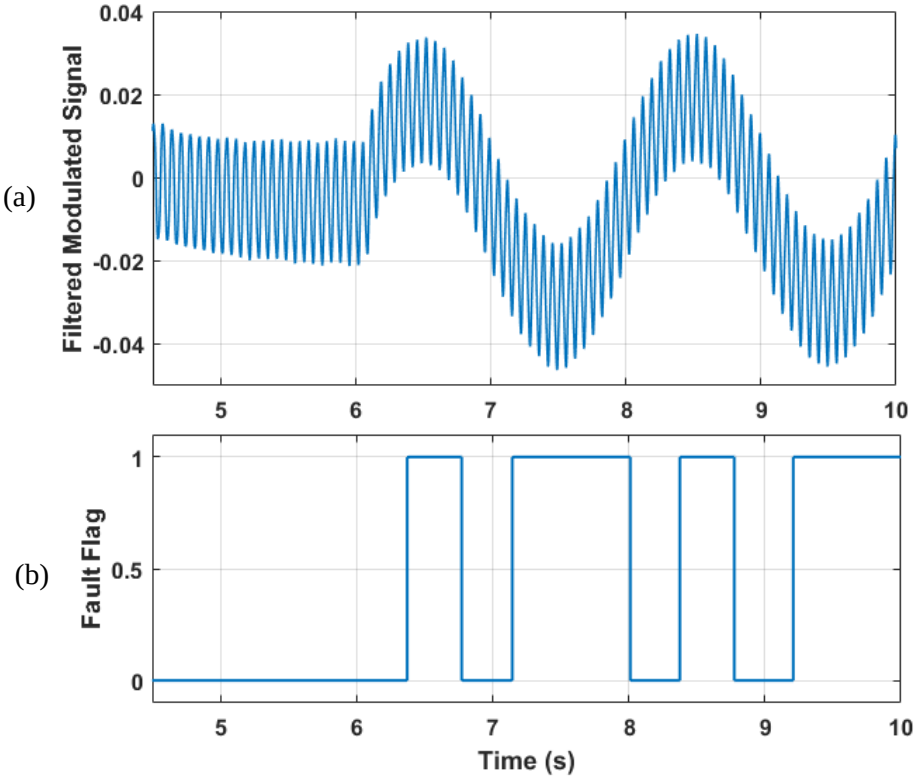

Fig. 11. The $y_{post}(t)$ signal for a DE/ME fault with a = .05A injected at 6.0525s (a). The subsequent fault flag signal (b) depicts the incorrect detection.

Table III. The minimum fault magnitude and corresponding fault sensitivity for each fault type

Fault Type	f_{fm} (Hz)	a_{min} (A)	Limit of fault sensitivity (%)
DE/ME	91.95	0.14	2.53
PSH	931.9	0.035	0.633
BRBF	65.85	0.17	3.08
SSWF	153.9	0.009	0.163
IRBF	205.9	0.0155	0.280
ORBF	157.9	0.015	0.271

Finally, Fig. 11 shows a scenario in which the magnitude a of the fault is too low relative to the fundamental magnitude, and the fault is improperly flagged. In this case, a DE/ME fault with a = 0.05A, or 0.9% of the fundamental, is injected at 6.0525s. Each time the slow fault-indicative component crosses zero, the ripples cause more than n_z zero crossings with less than time T_{zm} between them, so FC2 goes to 0 and the flag is removed. After running simulations for each fault with the same parameters as in Table II and d set to the average value of 0.5 Hz for all faults, the minimum fault magnitudes a_{min} that still result in complete correct fault detection are

summarized in Table III. Table III also shows the fault sensitivity of each fault type, as defined by a_{min}/A, where A = 5.526A in these simulations. The BRBF has the worst sensitivity at 3.08%, due to its low characteristic frequency, while the SSWF has the best fault sensitivity at an impressive 0.163%.

V. Conclusion

This paper described a novel time-domain method for fault detection and diagnosis that extend the recently developed adaptive modulation method to achieve shorter detection times. This is a simple method that requires minimal signal processing and no additional sensors, and it is capable of detecting all major types of induction motor faults concurrently and at a variety of motor speeds. The rapid detection characteristic of this method will prove invaluable for safety-critical and large manufacturing applications of induction motors. However, further developments are needed to optimize the model parameters and ensure the quickest detection time without compromising fault detection sensitivity and accuracy.

VI. References

[1] Y. Liu, A. Bazzi, B. Davis, "Adaptive modulation time-domain fault diagnosis of three-phase induction machines," presented at the *IEEE Int. Electric Machines and Drives Conf.*, Miami, FL, USA, May 21-24, 2017.

[2] Y. Liu, A.M. Bazzi, "A review and comparison of fault detection and diagnosis methods for squirrel-cage induction motors: state of the art," unpublished.

[3] S. Nandi, S. Ahmed, H.A. Toliyat, "Detection of rotor slot and other eccentricity related harmonics in a three phase induction motor with different rotor cages," *IEEE Trans. Energy Convers.*, vol. 16, no. 3, pp. 253-260, Sept. 2001.

[4] H. Fukui, Y. Mizuno, H. Nakamura, "Probabilistic diagnosis of short circuit fault in stator winding of motor based on feature distribution," in *IEEE Int. Symp. Electrical Insulation*, 2010, pp. 1-5.