Paper D

Improved Quadratic Time-frequency Distributions for Detecting Inter-turn Short Circuits of PMSM in transient States

Sveinung Attestog, Huynh Van Khang, Kjell G. Robbersmyr

This paper has been published as:

S. Attestog, H. V. Khang and K. G. Robbersmyr, "Improved Quadratic Time-frequency Distributions for Detecting Inter-turn Short Circuits of PMSMs in Transient States," 2020 International Conference on Electrical Machines (ICEM), Gothenburg, Sweden, 2020, pp. 1461-1467, doi: 10.1109/ICEM49940.2020.9271050.

D.1 Abstract

This paper aims to improve quadratic time-frequency distributions to adapt condition monitoring of electrical machines in transient states. Short-Time Fourier transform (STFT) has been a baseline signal processing technique for detecting fault characteristic frequencies. However, limits of window sizes due to loss of frequency- or time-resolution, make it hard to capture rapid changes in frequencies. Within this study, Choi-Williams and Wigner-Ville distributions are proposed to effectively detect peaks at characteristic frequencies while still maintaining low computation time. The improved quadratic time-frequency distributions allow for generating spectrograms of a longer lasting data signal and capturing multi-component signals with a better separation of the components than STFT. Further, the time resolution of the spectrograms generated by the proposed method is not affected by the window size. The effectiveness of the proposed methods is numerically verified from the data of an in-house test setup.

D.2 Introduction

Permanent magnet synchronous motors (PMSMs) are compact and highly efficient, making them attractive in electric powertrains for wind turbines and electric vehicles, which operate dynamically with variable speed and torque. Moreover, the powertrains are intensively exposed to mechanical-, and electrical stress in harsh environments and thermal cycling due to the dynamic operation. Consequently, detection and prevention of the faults in such powertrains are more important and challenging. In condition-based maintenance the machines are monitored over time, which allows to determine when the next maintenance is needed. Implemented correctly will reduce unexpected downtimes and costs.

Condition monitoring for electrical machines in dynamic operation requires the analysis of non-stationary signals. Short-time Fourier transform (STFT) is often used for this purpose. The resolution of the time-frequency representation, or spectrogram, is dependent on the window size. A smaller time window results in a high time resolution and fewer lines of resolution on the frequency axis. On the other hand, larger window sizes cause a lower time resolution. Over the years, multiple techniques have been developed, i.e., Wavelet transform [1] and quadratic time-frequency distribution (TFD), to address the computation burden and resolution. Quadratic TFD or Cohen Class function has been used in quantum physics and Heisenberg's uncertainty principle [2]. This family of functions includes Wigner-Ville (WVD), Choi-Williams (CWD), Zhao-Atlas-Marks (ZAM), and Rihaczek distributions. The difference among them is their kernel function. The advantage of these distributions is that they can capture the transient behavior of a signal, and time resolution is not affected by the length of the time sample. This property makes quadratic TFDs attractive in many applications besides fault detection of electrical motors, i.e., optical sensor, radar sensor, and wireless communication [3]–[5].

WVD is one of the oldest distributions and fast to compute. The demerit of this distribution is the presence of cross-terms. Additional peaks and patterns occur when applied to a signal with multiple components. Pseudo WVD or combining WVD with the

Gabor distribution can be a solution to address this challenge [6], [7]. Alternatively, an improved eigenvalue decomposition-based approach for reducing the cross-term was proposed in [8]. Minimizing the cross-term can also be done by changing the kernel function, giving a different Cohen class function. Towards condition monitoring for PMSMs, CWD was used for feature extraction in a demagnetization fault detection scheme [9]. The fault indicator was based on a box-counting fractal dimension. The spectrogram is divided into squares by a grid. The number of squares, where the signal is present, is counted. A demagnetization fault increases the number of boxes due to a more chaotic signal.

Although considerable researches have been devoted to enhance the accuracy of quadratic TFDs [6]–[13], limited research has focused on improving computation time of time-frequency representations or spectrograms [14]. The spectrograms were limited to shorter time samples. The largest time sample used in a bilinear TFD was found in [7], which was 2.7 s with sampling frequency 1.5 kHz. The increasing computation time with longer sample arrays restricts the application of quadratic TFDs in condition monitoring. This paper aims to propose an improved quadratic TFDs to make quadratic TFDs better suitable for condition monitoring of electrical machines in transient states while reducing the computational burden.

D.3 Mathematical background

D.3.1 Cohen Class Function

The general equation for Cohen class distribution function can be written as function of time t and frequency ω.

$$C(t,\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(\theta,\tau)\phi(\theta,\tau)e^{-j\theta t - j\tau\omega}d\theta d\tau \tag{D.1}$$

where $A(\theta,\tau)$ is the ambiguity function

$$A(\theta,\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R(t,\tau)e^{j\theta t}dt \tag{D.2}$$

$R(t,\tau)$ is the auto correlation function, which is defined as

$$R(t,\tau) = s^*(t-\frac{\tau}{2})s(t+\frac{\tau}{2}) \tag{D.3}$$

where τ is running time. The analytic signal $s(t)$ is defined as

$$s(t) = x(t) + jH(x(t)) \tag{D.4}$$

where the real part $x(t)$ is the original signal and the imaginary part $H(x(t))$ is the Hilbert transform of the original signal.

$$H(x(t)) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}d\tau \tag{D.5}$$

The integrals in (D.1) and (D.2) are Fourier and inverse Fourier transformations:

$$\mathcal{F}_{t\to\omega}(f(t)) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \tag{D.6}$$

$$\mathcal{F}^{-1}_{\omega\to t}(F(\omega)) = \int_{-\infty}^{\infty} F(\omega)e^{j\omega t}d\omega \tag{D.7}$$

The indexes $t \to \omega$ and $\omega \to t$ are added for clarifying what domains the data is switched between with the Fourier transformations. The general equation of Cohen class function, defined by (D.1) and (D.2), is rewritten to

$$C(t,\omega) = \mathcal{F}_{\tau\to\omega}\mathcal{F}_{\theta\to t}\left(\phi(\theta,\tau)\mathcal{F}^{-1}_{t\to\theta}\left(R(t,\tau)\right)\right) \tag{D.8}$$

The frequency domains θ and ω are different. The θ-domain is where the filtering by the kernel function is applied, while ω is the frequency domain of the resulting TFD.

The kernel function $\phi(\theta,\tau)$ is the main difference between the Cohen class functions. The simplest kernel functions is

$$\phi_{WVD}(\theta,\tau) = 1 \tag{D.9}$$

This gives the WVD, which can be written as

$$C(t,\omega) = \mathcal{F}_{\tau\to\omega}R(t,\tau) \tag{D.10}$$

The main problem of WVD is the cross-term when the signal $x(t)$ consists of multiple components. The most commonly used TFDs, that minimizes the cross-term, are CWD and ZAM, in which the kernel function is defined as:

$$\phi_{CWD}(\theta,\tau) = e^{-\frac{(\theta\tau)^2}{\sigma}} \tag{D.11}$$

and

$$\phi_{ZAM}(\theta,\tau) = \frac{\sin(\pi\theta\tau)}{\pi\theta\tau}e^{\frac{-2\pi\tau^2}{\sigma}} \tag{D.12}$$

Other notations of these kernel functions substitute σ with $\frac{1}{\alpha}$. Keeping σ is preferred, since reducing σ minimises the cross terms, while α has an inverse relation. Both $\phi_{CWD}(\theta,\tau)$ and $\phi_{ZAM}(\theta,\tau)$ become equivalent to $\phi_{WVD}(\theta,\tau)$ when σ goes toward ∞.

The integrals of the Cohen class distribution functions are defined from $-\infty$ to ∞. The next section will describe how to obtain the spectrograms for signals, that lasts for a finite period, with Fast Fourier Transform (FFT).

D.3.2 Discrete form

First apply the window function $d(n)$ to a measurement sample array $x(n)$. The selected window function is a hanning window.

$$d(n) = \frac{1}{2} + \frac{1}{2}\cos\left(2\pi\frac{n}{N}\right) \tag{D.13}$$

where N is the element size of $x(n)$ and n is the integer defined in the range $\left[-\frac{N}{2}, \frac{N}{2} - 1\right]$. The Hilbert transform is applied for obtaining the analytical signal $s(n)$ and a zero is appended,

$$s(n) = \begin{bmatrix} x(n) + jH\left(x(n)\right) \\ 0 \end{bmatrix}. \tag{D.14}$$

Then the elements of $s(n)$ need to be organized in the auto correlation matrix $R\left(n_t, n_\tau\right)$. The row vector n_τ and column vector n_t are defined as,

$$n_\tau = \begin{bmatrix} -\frac{N}{2} & \cdots & \frac{N}{2} - 1 \end{bmatrix} \tag{D.15}$$

and

$$n_t = \begin{bmatrix} -\frac{N}{2} & \cdots & \frac{N}{2} - 1 \end{bmatrix}^T \tag{D.16}$$

$R\left(n_t, n_\tau\right)$ is now written as,

$$R(n_t, n_\tau) = s^*(n_t - n_\tau) \odot s(n_t + n_\tau) \tag{D.17}$$

Note that the multiplications performed in (D.17) is a Hadamard product. About a half of the entries is not defined, because the sum or difference between n_t and n_τ is outside of the definition range of $x(n)$, $\left[-\frac{N}{2}, \frac{N}{2} - 1\right]$. These entries are set to 0. One way to solve this problem (in a code) is to set all entries, that are not defined in $R\left(n_t, n_\tau\right)$, equal to the last entry of $s(n)$, which is 0.

The ambiguity function is obtained by an inverse FFT from n_t-domain to n_θ-domain.

$$A\left(n_\theta, n_\tau\right) = \mathrm{FFT}^{-1}_{n_t \to n_\theta}\left(R(n_t, n_\tau)\right) \tag{D.18}$$

The characteristic function $M\left(n_\theta, n_\tau\right)$ equals the element wise product (Hadamard product) between the ambiguity function and the kernel function.

$$M\left(n_\theta, n_\tau\right) = \phi(n_\theta, n_\tau) \odot A(n_\theta, n_\tau) \tag{D.19}$$

The final TFD is obtained by a FFT from n_θ-domain back to n_t-domain and then by a FFT from n_τ-domain to n_ω-domain

$$C(n_t, n_\omega) = \mathrm{FFT}_{n_\tau \to n_\omega}\left(\mathrm{FFT}_{n_\theta \to n_t}\left(M\left(n_\theta, n_\tau\right)\right)\right) \tag{D.20}$$

Both window function and kernel function have been applied to the signal $x(n)$. Reducing σ for a CWD minimizes the cross-terms but spreads it across all frequencies e.g., raises the noise floor. It appears as vertical lines in a CWD with horizontal time axis and vertical frequency axis. The peaks of the frequency components are still present on top of the noise floor. This effect was filtered out by subtracting the hundredths smallest values of $C(n_t, n_\omega)$ for each time instant n_t.

D.4 Proposed Quadratic Time-Frequency Distribution

The essential parameters include the sampling frequency (f_s), time window (T), and overlap (OL), which are selected at 1 kHz, 1 s, and 20 %, respectively, in this work. Figure D.1 shows the flowchart of the proposed method with steps as follows.

1. Extract a sample from the original array from entry n_1 to n_2. The initial values of n_1 and n_2 are 1 and f_sT, respectively.

2. Execute the signal processing described in Section II on this sample and obtain the quadratic TFD.

3. Cut off the first 20 % and last 20 % of the spectrogram (OL). In the case, for the first second the spectrogram from 0.2 s to 0.8 s is extracted and put into the final plot.

4. Add the product of the OL and f_sT to n_1 to n_2 and repeat the process. In the second iteration, the next sample extracted is from 0.8 s to 1.8 s. The third interval will be from 1.6 s to 2.6 s.

5. Repeat until end of signal

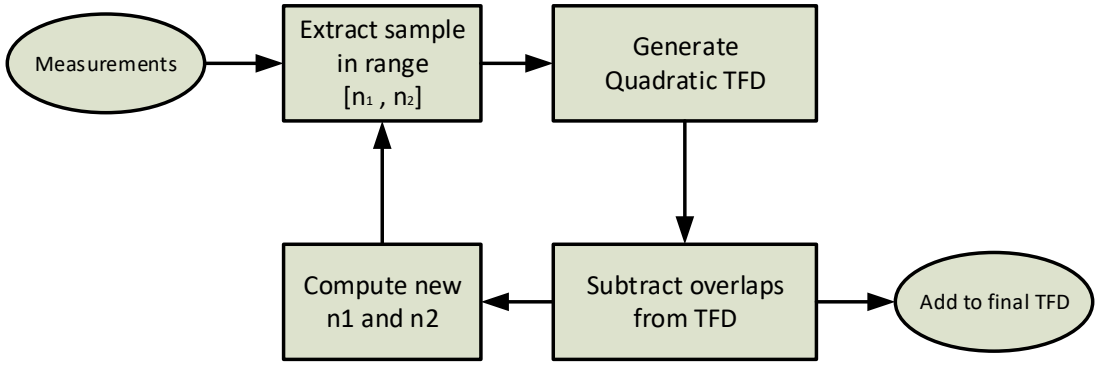


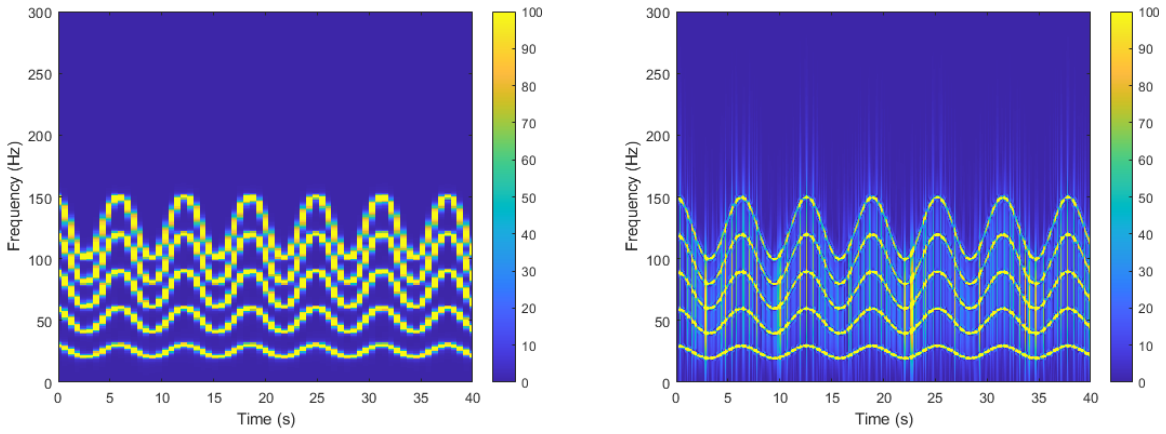
Figure D.1: Flowchart of the proposed quadratic TFD

D.5 Numerical test

In this section, the performance of the proposed quadratic TFDs is numerically investigated and compared with that of using STFT. In this numerical example, a signal with 5 frequency components used and was defined as:

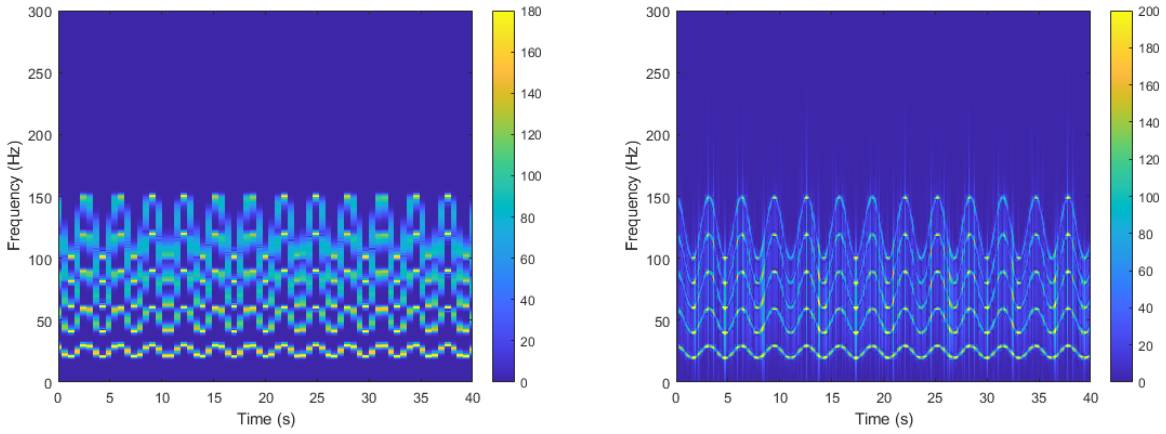
$$x(t) = \sum_{k=1}^{5} \sin\left(2\pi k\left(\frac{5}{c}\sin\left(ct\right) + 25t\right)\right). \tag{D.21}$$

The signal includes sinusoidal components with frequency modulation. The instantaneous frequency of the signal component, where $k = 1$ and $c = 1$ $\frac{\text{rad}}{\text{s}}$, is a sinusoidal function with amplitude of 1, frequency of $\frac{1}{2\pi}$ Hz and bias of 25 Hz. The constant c is set to 1 $\frac{\text{rad}}{\text{s}}$. Figs. D.2a and D.2b show the spectrograms generated with CWD and STFT, respectively. The window size was set 1 s with 20 % overlap. CWD was computed with $\sigma = 0.1$. Both algorithms capture all five signal components as described by (D.21). c was increased to 2 $\frac{\text{rad}}{\text{s}}$. The signal of (D.21) was separable in the CWD, but not separable in the spectrogram generated by STFT. (See Figure D.3a and Figure D.3b) The resolution improves when the window size is reduced to 0.1 s for STFT, but the signal component is



(a) (b)

Figure D.2: Numerical example - $c = 1$ and window 1 s - (a) STFT and (b) CWD



(a) (b)

Figure D.3: Numerical example - $c = 2$ and window 1 s - (a) STFT and (b) CWD

still hard to separate. The contours of the signal components in the CWD spectrogram seem to be thinner and have a smoother shape.

Decreasing window size to 0.1 s reduces the computation time for CWD from 16.5 s to 1.3 s, while the computation time for STFT is increased from 0.014 s to 0.040 s. The performance is considered acceptable for both methods since the total time for the measurement array was 120 s. The reduction in computation time for CWD is obtained since the CWD solves smaller matrices after each iteration. The algorithm solves a 100×100 matrix with a window size of 0.1 s and a 1000×1000 matrix with a window size of 1 s. The sampling rate is 1 kHz. Both CWD and STFT need to perform more iterations due to a smaller window. Another interesting finding is that increasing the window size for CWD does not affect the time resolution of the spectrogram. Changes in frequency can still be captured by CWD, while all the signal components will get merged in the STFT at too larger window sizes. However, the window size only needs to be sufficiently large for the CWD. Smaller window size will give problems with lines of resolution on the frequency axis. Note that a larger window size increases the computation time.

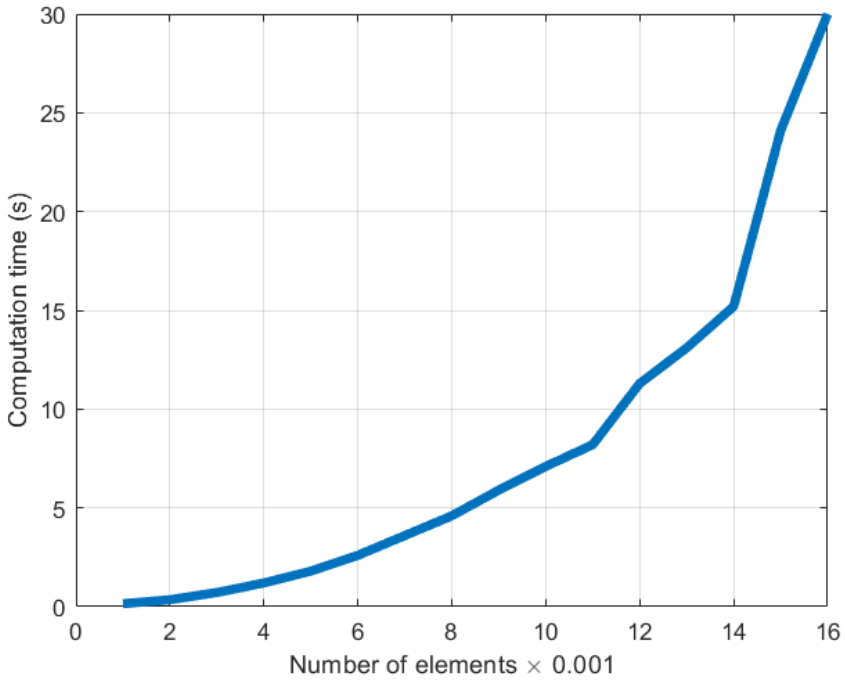


Figure D.4: Computation time versus number of elements in array - CWD

Figure D.4 shows the computation time for executing one iteration of the proposed algorithm presented versus the element size of the input sample. The overlap is set to 0 %. If the input array has an element size of 10000, it will take 7.5 s to solve. An array of 10000 elements can have a total time of 1 s with sample rate of 10 kHz or 10 s with sample rate of 1 kHz. The second sample would be possible in condition monitoring, because the total time would is larger than the computation time.

The replacement of values in the matrices of $n_t - n_\tau$ and $n_t + n_\tau$ is one of the time consuming operations. This is solved with if-statements, where half of the entries end up as 0 in $R(n_t, n_\tau)$. One possibility for reducing computation complexity is to avoid computing parts of the kernel function and auto-correlation matrix because those entries have an insignificant effect on the final TFD. This is explored in detail in [14]. The complexity of FFT (and inverse FFT) applied to a vector with N elements is $O(N \log(N))$. Both $R(n_t, n_\tau)$ and $M(n_\theta, n_\tau)$ are square matrices with N rows and N columns. Inverse FFT is executed once, and FFT is executed twice on $N \times N$ matrices. Therefore, the computational complicity of all these transformations is $O(3N^2 \log(N))$, which explains why the computational complexity increase exponentially in Figure D.4. The performance of the algorithm can can be improved with faster computer or a more efficient code.

D.6 Experimental results and Discussions

The experimental setup consists of two 2.5 kW, 16 poles PMSMs as shown in Figure D.5, in which one operates as a motor while another motor operates as a generator coupled to a resistive load. An inter-turn winding short circuit fault with 10 % severity is seeded

in the motor. The setup is tested on two different speed profiles in both healthy and faulty conditions. The first one is a multi-step speed-profile between standstill to 375 rpm and back down to a full stop. The second speed-profile is an unpredictable speed-profile where the speed increases and decreases around 250 rpm in an irregular manner. The experimental setup is described in detail in [15].

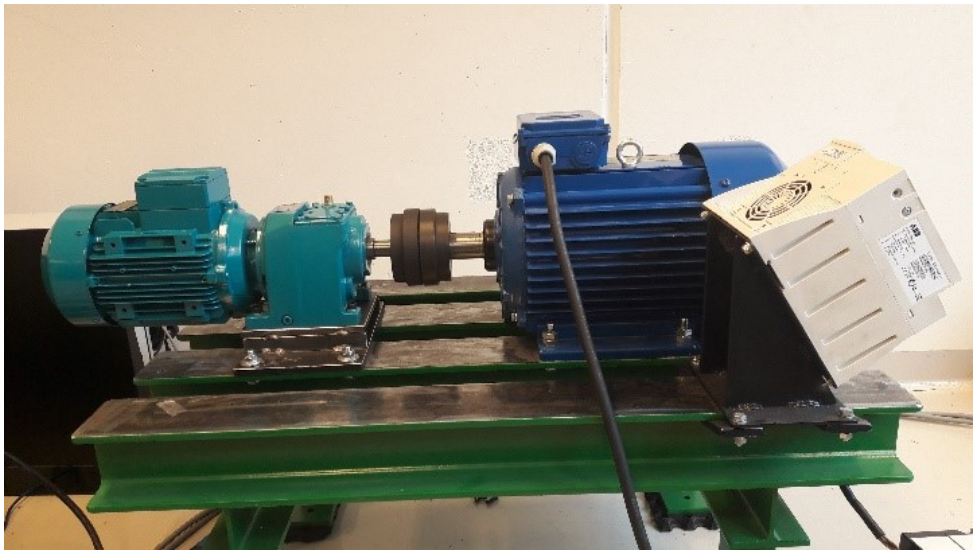
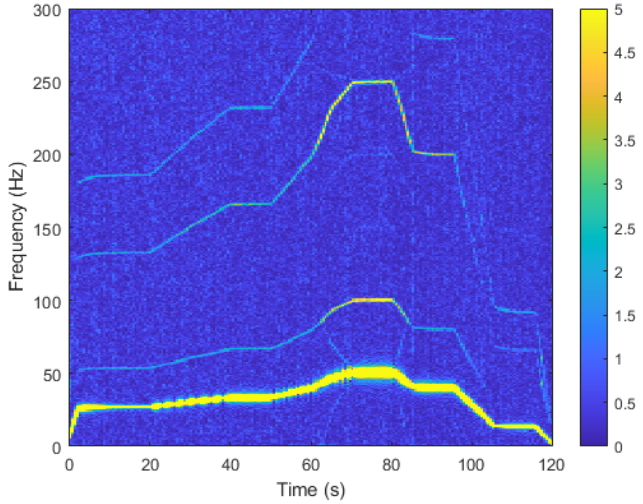


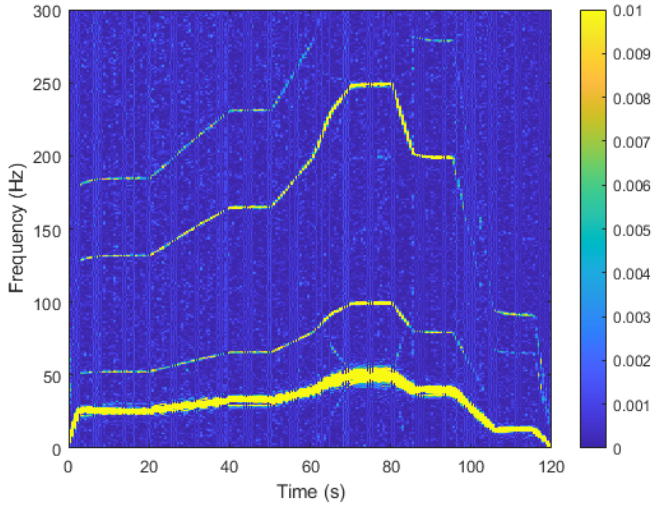
Figure D.5: The in-house experimental setup.

Figure D.6a shows the spectrogram generated by STFT for a phase current of the PMSM in the healthy case. Sampling rate and window size are 1 kHz and 1 s, respectively. The hanning window function was applied to the window with an overlap of 20 %. STFT is used as a benchmark to compare with the spectrograms generated by CWD and WVD. Figure D.6b and D.6c show the spectrograms generated from CWD and WVD. The kernel function for CWD has $\sigma = 10^{-7}$. The harmonic peaks were captured and proved that CWD is capable of capturing harmonic peaks. The fundamental frequency is much more dominant in CWD as compared to the STFT. This may be due to the extremely small σ. One attempt to solve this issue was done by applying an element-wise square root of the CWD, but this makes the spectrograms noisier. The additional filtering by subtracting the hundredths smallest value for each instant in time improved the quality of the spectrogram. The WVD-based spectrograms have additional peaks, but other peaks are not visible on the spectrograms. This is caused by the cross-term and the limits on the color axis.

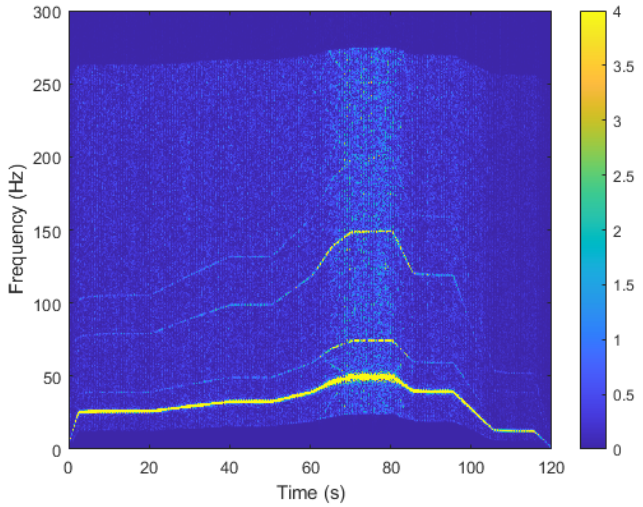
Figure D.7a, D.7b and D.7c show the spectrograms obtained by STFT, CWD, and WVD of the phase current for the same PMSM operating under an inter-turn short circuit. Both STFT and CWD was able to pick up the second harmonic, which is an indicator for electrical faults. WVD seems to pick up all the integer harmonics. Similar results are shown in Figure D.8a, D.8b and D.8c, which show the spectrograms for the second speed profile generated by STFT, CWD, and WVD, respectively. These experimental results proved that CWD can pick up the same harmonic peaks as STFT. The fundamental component is much more dominant in the CWD. It was discussed earlier in the numerical example, how the quadratic TFDs would perform better with more rapid speed changes.



(a)

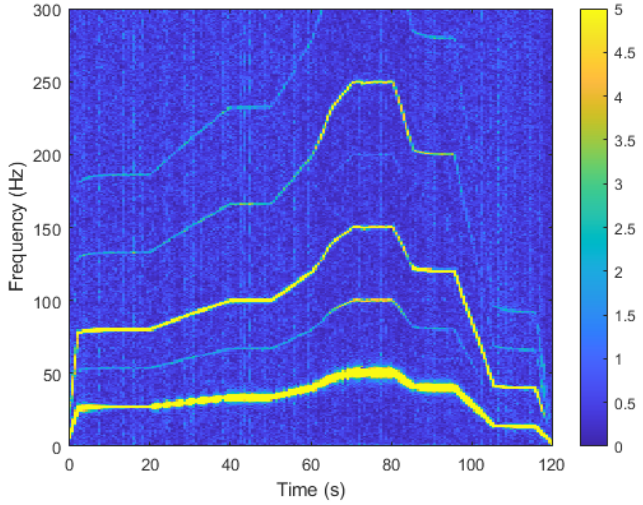


(b)

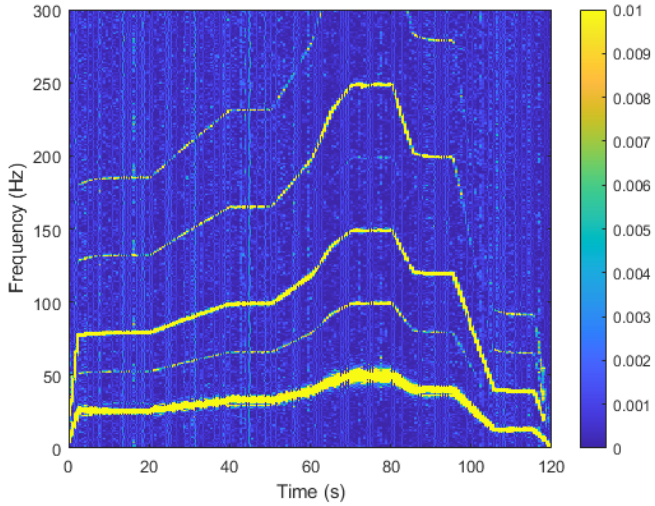


(c)

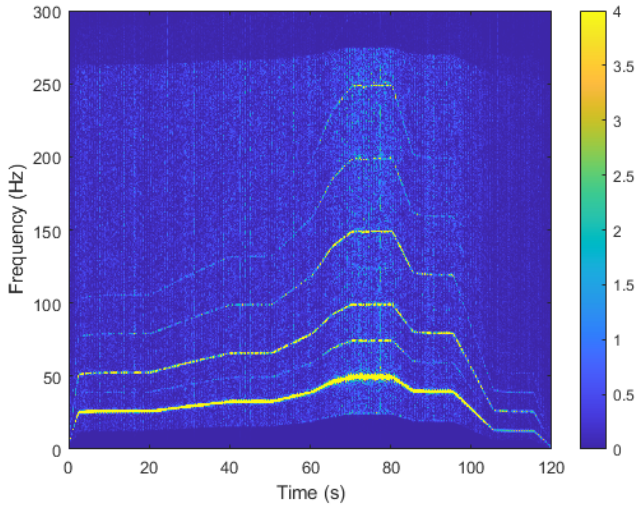
Figure D.6: Spectrograms of phase current of healthy PMSM operating with the first speed profile, (a) STFT, (b) CWD and (c) WVD



(a)

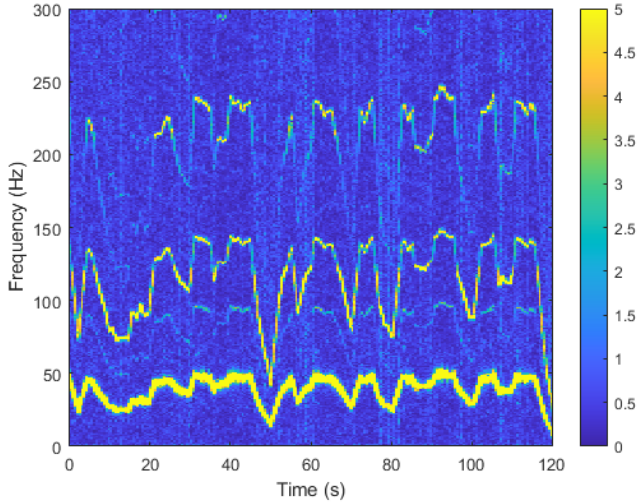


(b)

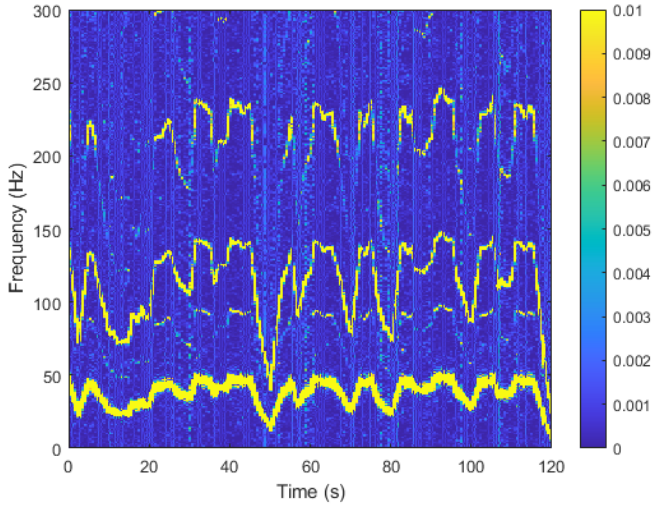


(c)

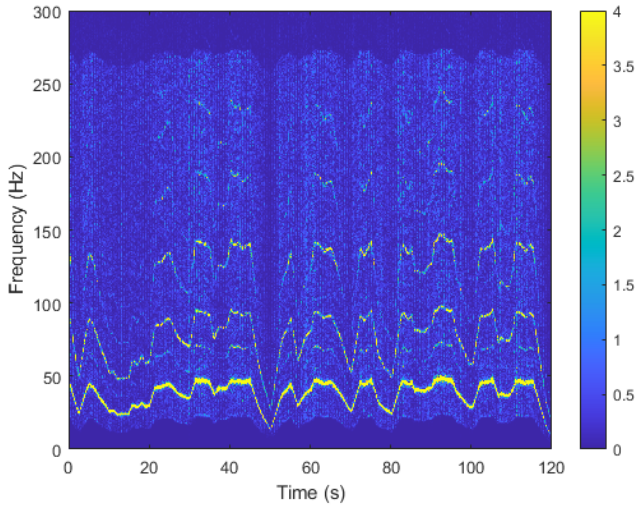
Figure D.7: Spectrograms of phase current of PMSM with 10 % ITSC operating with the first speed profile, (a) STFT, (b) CWD and (c) WVD



(a)



(b)



(c)

Figure D.8: Spectrograms of phase current of PMSM with 10 % ITSC operating with the second speed profile, (a) STFT, (b) CWD and (c) WVD

The second and third harmonics have merged in the first 10 s and around 50 s on the spectrograms using STFT and CWD as shown in Figure D.8a and Figure D.8b, respectively. The reason for this issue is that they are separable in the WVD (See Figure D.8c). Multiplying the fundamental frequency with integers would also lead to the same conclusion.

The spectrograms computed from the experimental data shows that the quadratic TFD can capture the signal, but the current set-up does not give any additional information compared to STFT. But further analysis shows that the signal components in the CWD have sharper peaks. The spectrogram generated from the second speed profile was divided into three frequency bands as listed in Table I. The first, third and fifth harmonics are present and dominant in these bands. The time average kurtosis comparison between CWD and STFT is presented in Table I. CWD have a larger kurtosis in all three frequency bands, which indicates that it got higher and sharper central peaks.

Table D.1: Average kurtosis over frequency bands in the spectrgram generated by STFT and CWD

Frequency band	CWD	STFT
0 Hz to 50 Hz	414.0	20.5
50 Hz to 150 Hz	9.6	2.6
150 Hz to 300 Hz	4.0	2.1

D.7 Conclusion

In this paper, we have proposed a solution to improve quadratic TFD for analyzing non-stationary signals. The proposed method allows for generating spectrograms with CWD and WVD for longer-lasting signals, as compared to the existing methods. The numerical test showed that CWD could generate spectrograms with better separation of multiple components than STFT if the frequency modulation is too large. Within the framework, the time resolution is unaffected by window size, but frequency resolution sets a minimum limit on the window. Larger windows increase the computation time and set a maximum limit. The test on the experimental data proved that the proposed method could capture the signal components on data with noise. The time average kurtosis for generating spectrograms by CWD was larger than that of STFT for the first, third and fifth harmonics of the signal. This indicates that CWD generates spectrograms with sharper peaks.

Future work will include trying out other distribution functions that will overcome the problems of the CWD, and make the fundamental frequency less dominant. The WVD has a problem with cross-terms, but is much faster than other Cohen class functions. Alternatively, a pseudo WVD could be implemented instead, which would reduce cross-terms and maintain high computation speed and resolution.

References

[1] M. Heydarzadeh, M. Zafarani, M. Nourani, and B. Akin, "A wavelet-based fault diagnosis approach for permanent magnet synchronous motors," *IEEE Transactions on Energy Conversion*, vol. 34, no. 2, pp. 761–772, Jun. 2019.

[2] L. Cohen, "Time-frequency distributions-a review," *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941–981, Jul. 1989.

[3] Y. Yu, L. Luo, B. Li, K. Soga, and J. Yan, "Quadratic time-frequency transforms-based brillouin optical time-domain reflectometry," *IEEE Sensors Journal*, vol. 17, no. 20, pp. 6622–6626, Aug. 2017.

[4] J. Gao, L. Shen, S. W. F. Ye, and R. Zhangn, "Multi-feature radar signal modulation recognition based on improved PSO algorithm," *The Journal of Engineering*, vol. 2019, no. 19, pp. 5588–5592, Jun. 2019.

[5] T. Wang, J. Liang, and X. Liu, "Soil Moisture Retrieval Algorithm Based on TFA and CNN," *IEEE Access*, vol. 7, pp. 597–604, 2019.

[6] Y. S. Shin and J. Jeon, "Pseudo Wigner-Ville Time-Frequency Distribution and its Application to machinery condition monitoring," *Shock and Vibration*, vol. 1, pp. 65–76, 1993.

[7] A. L. Martinez-Herrera, L. M. Ledesma-Carrillo, M. Lopez-Ramirez, S. Salazar-Colores, E. Cabal-Yepez, and A. Garcia-Perez, "Gabor and the Wigner-Ville transforms for broken rotor bars detection in induction motors," in *2014 International Conference on Electronics, Communications and Computers (CONIELECOMP), Cholula*, May 2014, pp. 83–87.

[8] R. Sharma and R. B. Pachori, "Improved Eigenvalue Decomposition-Based Approach for Reducing Cross-Terms in Wigner–Ville Distribution," *Circuits, Systems, and Signal Processing*, vol. 37, no. 8, pp. 3330–3350, 2018.

[9] M. Delgado Prieto, A. Garcia Espinosa, J.-R. Riba Ruiz, J. C. Urresty, and J. A. Ortega, "Feature extraction of demagnetization faults in permanent-magnet synchronous motors based on box-counting fractal dimension," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1594–1605, May 2011.

[10] S. Rajagopalan, J. A. Restrepo, J. M. Aller, T. G. Habetler, and R. G. Harley, "Nonstationary motor fault detection using recent quadratic time–frequency representations," *IEEE Transactions on Industry Applications*, vol. 44, no. 3, pp. 735–744, May 2008.

[11] J. Zhang, H. Gao, Q. Liu, and C. Grebe, "A new real-time signal processing approach for frequency-varying machinery," *Journal of Vibration and Control*, vol. 24, no. 11, pp. 2359–2368, 2018.

[12] Y. Guan, M. Liang, and D. S. Necsulescu, "Velocity synchronous bilinear distribution for planetary gearbox fault diagnosis under non-stationary conditions," *Journal of Sound and Vibration*, vol. 443, pp. 212–229, 2019.

[13] W. Liu and X. Guo, "Detection of transient power quality disturbances based EMD combined with Choi-Williams distribution," in *2012 IEEE International Conference on Automation and Logistics*, pp. 588–591.

[14] B. Hollinger Kenneth, *Code optimization for the Choi-Williams distribution for ELINT applications*, Dec. 2009. [Online]. Available: `https://calhoun.nps.edu/handle/10945/4422`.

[15] J. S. L. Senanayaka, V. K. Huynh, and K. G. Robbersmyr, "Fault detection and classification of permanent magnet synchronous motor in variable load and speed conditions using order tracking and machine learning," *Journal of Physics: Conference Series*, vol. 1037, no. 3, 2018.