Research Article

Adaptive Optimal Kernel Smooth-Windowed Wigner-Ville Distribution for Digital Communication Signal

Jo Lynn Tan and Ahmad Zuri bin Sha'ameri

Department of Microelectronic and Computer Engineering, Universiti Teknologi Malaysia, Skudai 81310, Johor, Malaysia

Correspondence should be addressed to Jo Lynn Tan, tjolynn82@yahoo.co.uk

Time-frequency distributions (TFDs) are powerful tools to represent the energy content of time-varying signal in both time and frequency domains simultaneously but they suffer from interference due to cross-terms. Various methods have been described to remove these cross-terms and they are typically signal-dependent. Thus, there is no single TFD with a fixed window or kernel that can produce accurate time-frequency representation (TFR) for all types of signals. In this paper, a globally adaptive optimal kernel smooth-windowed Wigner-Ville distribution (AOK-SWWVD) is designed for digital modulation signals such as ASK, FSK, and *M-ary* FSK, where its separable kernel is determined automatically from the input signal, without prior knowledge of the signal. This optimum kernel is capable of removing the cross-terms and maintaining accurate time-frequency representation at SNR as low as 0 dB. It is shown that this system is comparable to the system with prior knowledge of the signal.

1. INTRODUCTION

Bilinear time-frequency analysis has been widely used to analyze time-varying signals such as in speech, music and other acoustical signals, sonar, radar, geophysics, and biological signals. However, a major drawback of this method is the presence of cross-terms in the time-frequency representations (TFRs) [1]. If the cross-terms are not minimized in the time-frequency distribution (TFD), they will reduce the autoterms resolution and make interpretation of the true signal characteristics difficult [2]. To overcome this, most of the TFDs employ some kind of smoothing kernel, window, or filter [3]. Smoothing, however, causes the autoterms to be smeared and as a result the TFR losses its concentration [4]. For signal analysis and classification, an optimal distribution should have reasonable cross-terms suppression and minimal smearing of the autoterms. Previous works have shown that the optimal kernel is signal-dependant [2, 3, 5]. Generally, there is no single TFD with a fixed window or kernel which would perform well for all signals. A kernel might perform very well for a certain class of signals but is not optimal for other types of signals. For optimal TFR, the selection of appropriate kernel requires prior knowledge of the signal components under observation, which are usually not available in many applications. With this in mind, we aim to design an optimal kernel that will work in noncooperative environment, where signals are unknown in nature.

Adaptive kernel, which is capable to change according to the signal of interest, will be able to give optimal TFR for a substantially wide range of signal types. Several researchers have developed the adaptive kernel TFRs, which are optimized either globally and applied to the entire signal [5, 6], or optimized at every time instant or every frequency interval [2–4]. The globally adapted kernel is inappropriate for signals whose time-frequency behavior changes with time or frequency such as in multicomponent signals because the kernel will not be able to adapt with the changes within the evaluation period. Whenever the signal parameter changes, it will fail to produce the optimal TFR. Locally adapted kernel, on the other hand, will be able to detect the changes and optimize accordingly but it requires extensive [2] or repeated computation algorithms [4]. Due to the computational complexity, some of these methods are not suitable for real-time analysis [2, 4] unless they are optimized [3]. Most of the researches in this area focus mainly on linear FM [2, 3, 5, 6]

and biological signals [7, 8]. Not much attention has been given to digital communication signals.

This paper suggests a globally adaptive optimal kernel smooth-windowed Wigner-Ville distribution (AOK-SWWVD) for digital modulation signal such as ASK, FSK, and *M-ary* FSK. These signals are time-varying signals which frequencies vary with time but are time-invariant in their modulation parameters such as symbol rate and frequency deviation. The optimal kernel depends only on the modulation parameters of the signals of interest, which are assumed invariant throughout the evaluation period. Thus, a globally adapted kernel is used to avoid the unnecessary computations in locally adapted kernel. Optimal kernel in our context is a kernel which gives a TFR with minimal smearing of the autoterms and strong suppression of the cross-terms components.

This correspondence is organized as follows. In Section 2, we give a summary of signals that are used for the evaluation in this paper. A brief discussion on bilinear time-frequency distribution is given in Section 3. In Section 4, the general equations of the bilinear product in time-lag domain for both autoterms and cross-terms are derived. The kernel parameters are then determined mathematically for the FSK and ASK signals. A guideline on how to determine the kernel parameters for optimal TFR is given. Based on these guidelines for optimal kernel design, an adaptive system which requires no prior knowledge of the signal is designed in Section 5. Section 6 shows the performance comparison between this adaptive system and an optimal system where its kernel is mathematically designed based on prior knowledge of the signal. They are compared in terms of main-lobe width (MLW), peak-to-side lobe ratio (PSLR), bias in symbol-duration (SDB), and signal-to-cross terms ratio (SCR). Conclusions are given in Section 7.

2. SIGNAL MODELS

Types of digital modulation which are considered in this paper are ASK, FSK, and *M-ary* FSK. These signals are commonly used in the digital communication. Consider an arbitrary digital communication signal, formed as a sum of N short-duration complex exponential signals, given as follows:

$$z(t) = \sum_{k=1}^{N} A_k \exp\left(j2\pi f_k(t - (k-1)T_b + \varphi)\right) \\ \times \Pi(t - (k-1)T_b), \tag{1}$$

where k is the binary sequence number starting with one, A_k is the amplitude, f_k is the subcarrier frequency, φ is the phase, and T_b is the symbol duration of the signal. ASK signal has constant frequency f_k and phase φ, but its amplitude A_k changes according to the symbol sequence transmitted, $A_k = 1$ when symbol "1" and $A_k = 0$ when symbol "0."

FSK and *M-ary* FSK signals have constant amplitude A_k and phase φ, but varying frequency f_k according to the symbol sequence sent. f_k is the subcarrier frequency at kth symbol for FSK and *M-ary* FSK. For FSK signal, $f_k = f_0$

when symbol "0" and $f_k = f_1$ when symbol "1." For *M-ary* FSK, f_k is set according to the combination of bits in a symbol. For all signals, the box function is defined as

$$\Pi(t) = 1, \quad \text{for } 0 \le t \le T_b, \\ = 0, \quad \text{elsewhere.} \tag{2}$$

The signal parameters of the signals used in this paper are given as follows:

(1) FSK1: $f_0 = 2125$ Hz, $f_1 = 2295$ Hz, $T_b = 20$ ms, $\varphi = 0$;

(2) FSK2: $f_0 = 2125$ Hz, $f_1 = 2295$ Hz, $T_b = 13.33$ ms, $\varphi = 0$;

(3) FSK3: $f_0 = 2125$ Hz, $f_1 = 2295$ Hz, $T_b = 10$ ms, $\varphi = 0$;

(4) FSK4: $f_0 = 2125$ Hz, $f_1 = 2295$ Hz, $T_b = 8$ ms, $\varphi = 0$;

(5) ASK: $f_0 = 2000$ Hz, $T_b = 10$ ms, $\varphi = 0$;

(6) 8FSK: 600 Hz $\le f_k \le 2000$ Hz, $T_b = 20$ ms, $f_{\text{dev}} = 200$ Hz, $\varphi = 0$;

(7) 16FSK: 400 Hz $\le f_k \le 3400$ Hz, $T_b = 20$ ms, $f_{\text{dev}} = 200$ Hz, $\varphi = 0$.

3. BILINEAR TIME-FREQUENCY DISTRIBUTION

The bilinear formulation for time-frequency distributions [9] is given as

$$\rho_z(t, f) = \int_{-\infty}^{\infty} G(t, \tau) \underset{(t)}{*} K_z(t, \tau) \exp(-j2\pi f\tau) d\tau, \tag{3}$$

where $G(t, \tau)$ is the time-lag kernel function and $K_z(t, \tau)$ is the bilinear product. The bilinear product is further defined as

$$K_z(t, \tau) = z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right), \tag{4}$$

where $z(t)$ is the analytic signal of interest. In this paper, we use a separable kernel which is separated in time and lag such that

$$G(t, \tau) = H(t)w(\tau), \tag{5}$$

where $H(t)$ is the time-smooth (TS) function and $w(\tau)$ is the lag-window function. The separable kernel smooth-window Wigner-Ville distribution (SWWVD) is given as

$$\rho_{z,\text{SWWVD}}(t, f) = \int_{-\infty}^{\infty} H(t) \underset{(t)}{*} K_z(t, \tau) w(\tau) \exp(-j2\pi f\tau) d\tau. \tag{6}$$

Any function similar to the popular window functions used in filter design or spectrum analysis or pulse shaping functions in digital communications can be used as the lag-window and the TS function. For a distribution with reduced cross-terms, the kernel used should be a low-pass window in lag domain and low-pass filter in the Doppler domain

(*Doppler domain*, v *is the Fourier Transform pair of time domain*, t *as the frequency domain*, f *is the Fourier Transform pair of lag domain* τ) as the autoterms are concentrated around the origin and the lag and Doppler coordinate axes [10, 11]. We use Hamming window as the lag-window and raised cosine pulse as the TS function. Hamming window is given as

$$w(\tau) = 0.54 + 0.46 \cos \frac{\pi \tau}{T_g}, \quad |\tau| \le T_g. \tag{7}$$

This lag window has the cutoff lag at

$$\tau_c = T_g. \tag{8}$$

Raised-cosine pulse is given as

$$H(t) = 1 + \cos\left(\frac{\pi t}{T_{sm}}\right), \quad 0 \le t \le T_{sm,}$$
$$= 0, \quad \text{elsewhere.} \tag{9}$$

The Doppler representation of this TS function obtained from the Fourier transform with respect to time for $H(t)$ is

$$h(v) = \frac{\sin(\pi v T_{sm})}{\pi v T_{sm}} + \frac{1}{2} \frac{\sin(\pi(v - 1/2T_{sm}))}{\pi(v - 1/2T_{sm})}$$
$$+ \frac{1}{2} \frac{\sin(\pi(v + 1/2T_{sm}))}{\pi(v + 1/2T_{sm})}. \tag{10}$$

It is a low-pass filter in the Doppler domain, where the cutoff Doppler is

$$v_c = \frac{3}{2T_{sm}}. \tag{11}$$

4.   TIME-LAG REPRESENTATION

This section describes the general bilinear product of the signals of interest in the time-lag domain and how the information is used to determine the kernel parameters.

4.1.   *Bilinear product of digital modulation signals*

For an arbitrary digital modulation signal, the time-lag representations of the bilinear product defined in terms of the autoterms and cross-terms are given as follows. The derivation from (12) to (17) is given in the appendix:

$$K_z(t, \tau) = K_{z,\text{auto}}(t, \tau) + K_{z,\text{cross}}(t, \tau), \tag{12}$$

$$K_{z,\text{auto}}(t, \tau) = \sum_{k=1}^{N} K_{z,k,k}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right), \tag{13}$$

$$K_{z,\text{cross}}(t, \tau)$$
$$= \sum_{\substack{k=1, l=1 \\ k \neq l}}^{N} \sum K_{z,k,l}\left(t - \frac{(k+l-1)T_b}{2}, \tau - (k-l)T_b\right), \tag{14}$$

where k and l represent the sequence of symbol present in the received signal. Both the kth and lth autoterms and cross-terms components in (13) and (14) are further defined as

$$K_{z,k,k}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right)$$
$$= |A_k|^2 \exp\left(j2\pi f_k \tau\right) K_{\Pi}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right), \tag{15}$$

$$K_{z,k,l}\left(t - \left(\frac{(k+l-1)T_b}{2}\right), \tau - (k-l)T_b\right)$$
$$= A_k A_l^* \exp\left(j2\pi((k-1)f_k - (l-1)f_l)T_b\right)$$
$$\times \exp\left(j2\pi\left(\frac{(f_k + f_l)}{2}\right)\tau\right)$$
$$\times \exp\left(j2\pi(f_l - f_k)t\right) K_{\Pi}\left(t - \left(\frac{(k+l-1)T_b}{2}\right),\right.$$
$$\left. \tau - (k-l)T_b\right), \tag{16}$$

where f_k and f_l represent the frequency of the symbol and A_k and A_l represent the amplitude of the symbol. The bilinear product of the box function $\Pi(t)$ defined in (2) is defined as

$$K_{\Pi}\left(t - \left(\frac{(k+l-1)T_b}{2}\right), \tau - (k-l)T_b\right)$$
$$= \Pi\left(t - kT_b + \frac{\tau}{2}\right)\Pi\left(t - lT_b - \frac{\tau}{2}\right). \tag{17}$$

For a given kth auto-term, the single-lag component with the frequency f_k lies along the time axis at lag $\tau = 0$. On the other hand, the cross-term between kth and lth symbol has Doppler frequency component at $v = (f_l - f_k)$ and lag-frequency component at $f = (f_k + f_l)/2$ and is located at lag $|\tau| > 0$. This is consistent with the findings by various researchers [2–6, 10, 11] which state that autoterms are concentrated along the axis while the cross-terms are located away from the axis. By choosing appropriate parameters for the separable kernels, the autoterms can be preserved while the cross-terms are suppressed. The cross-terms can be suppressed by using low-pass filter and low-pass window. Suitable length of TS function $H(t)$ removes the Doppler frequency, v components, while appropriate window width of lag-window, $w(\tau)$, removes cross-terms that lie at lag $|\tau| > 0$.

4.2.   *Bilinear product of FSK/M-ary FSK signal*

For simplicity, we will first evaluate FSK signal of 4 symbols length in the time-lag domain. The same argument can be used for signal of other symbol length and for *M*-ary FSK signal. The time-lag representation for the FSK will be represented based on a binary sequence of "1101" and the modulation parameters defined in Section 2. Discussion will

be on selected autoterms and cross-terms components. For this signal, $f_k = f_l$ for $k = 1, 2, 4$ and $f_k = f_0$ for $k = 3$. Based on (15), the auto-term at $k = 2$ is

$$K_{z,2,2}\left(t - \frac{3T_b}{2}, \tau\right) = \exp\left(j2\pi f_1 \tau\right) K_{\Pi}\left(t - \frac{3T_b}{2}, \tau\right). \quad (18)$$

This function is centered at time $t = 3T_b/2$ and lag $\tau = 0$. Autoterms are generated by the autocorrelation of the same symbol. On the other hand, cross-terms are generated by the correlation of different symbols. The cross-term between different symbols that have the same frequency but fall at different time instants can be seen at $k = 1$ and $l = 4$, which refers to the interaction between the 1st and the 4th symbols. From (16), the cross-term is expressed as

$$K_{z,1,4}(t - 2T_b, \tau + 3T_b)$$
$$= \exp\left(j2\pi f_1(\tau - 3T_b)\right) K_{\Pi}(t - 2T_b, \tau + 3T_b). \quad (19)$$

This cross-term is centered at $t = 2T_b$ and $\tau = -3T_b/2$. The cross-term between symbols that have different frequency can be seen at $k = 2$ and $l = 3$, which refers to the interaction between the 2nd and the 3rd symbols. From (16), this is expressed as

$$K_{z,2,3}(t - 2T_b, \tau + T_b)$$
$$= \exp\left(j2\pi(f_1 - 2f_0)T_b\right) \exp\left(j2\pi\left(\frac{(f_1 + f_0)\tau}{2}\right)\right)$$
$$\times \exp\left(j2\pi(f_0 - f_1)t\right) K_{\Pi}(t - 2T_b, \tau + T_b). \quad (20)$$

This cross-term is centered at time $t = 2T_b$ and lag $\tau = -T_b$ with Doppler-frequency component of $v = (f_0 - f_1)$ and lag-frequency component of $f = (f_0 + f_1)/2$.

All autoterms and cross-terms of the bilinear product for the FSK signal are shown in Figure 1. Autoterms are lightly dotted while the cross-terms are densely dotted. From Figure 1, we can see that, in general, the autoterms lie along the time axis and centered at lag, $\tau = 0$, while the cross-terms are elsewhere. To preserve the concentration of the autoterms while removing cross-terms, a lag-window should cover all the autoterms while removing the cross-terms as much as possible. The lag-window width, T_g, can be set such that

$$|T_g| \le T_b. \quad (21)$$

By setting this limit, the whole autoterms, which are along the time axis, can be preserved. However, unavoidably, part of the cross-terms such as at $k = 2$, $l = 3$ and $k = 3$, $l = 4$ is also preserved due to their adjacency to the autoterms as shown in Figure 1. These adjacent cross-terms contribute as interference if they have nonzero Doppler frequency. A smaller lag-window width could remove more of the adjacent cross-terms but at a price of reducing the autoterms concentration and causes smearing in the TFD.

By not minimizing the lag window further, a TS function is included in the SWWVD. The TS function acts like a low-pass filter in the Doppler frequency, v domain, as shown in (10). It removes the Doppler-frequency components of the remaining cross-terms which cannot be removed by the lag-window due to their adjacency to the autoterms. The smoothed bilinear product, $R_{z,sm}(t, \tau)$, is a convolution between the TS function and the bilinear product of the signal which relates to (6):

$$R_{z,sm}(t, \tau) = H(t) \underset{(t)}{*} K_z(t, \tau). \quad (22)$$

The smoothed bilinear product of the autoterms is given as

$$R_{z,sm,k,k}(t, \tau)$$
$$= H(t) \underset{(t)}{*} \exp\left(j2\pi f_k \tau\right) K_{\Pi}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right)$$
$$= h(v)|_{v=0} \exp\left(j2\pi f_k \tau\right) K_{\Pi}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right)$$
$$= h(0) \exp\left(j2\pi f_k \tau\right) K_{\Pi}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right). \quad (23)$$

Since we want to preserve the autoterms, the cutoff Doppler-frequency is set as $v_c > 0$. The smoothed bilinear product of the cross-terms is given as

$$R_{z,sm,k,l}(t, \tau)$$
$$= H(t) \underset{(t)}{*} e^{j2\pi(f_k - f_l)t} e^{j2\pi(f_1 + f_0)\tau/2} e^{j2\pi((l-1)f_l - (k-1)f_k)T_b}$$
$$\times K_{\Pi}\left(t - \frac{(k+l-1)T_b}{2}, \tau - (k-l)T_b\right)$$
$$= h(v)|_{v=f_k - f_l} e^{j2\pi(f_k - f_l)t} e^{j2\pi(f_1 + f_0)\tau/2} e^{j2\pi((l-1)f_l - (k-1)f_k)T_b}$$
$$\times K_{\Pi}\left(t - \frac{(k+l-1)T_b}{2}, \tau - (k-l)T_b\right)$$
$$= h(f_k - f_l) e^{j2\pi(f_k - f_l)t} e^{j2\pi(f_1 + f_0)\tau/2} e^{j2\pi((l-1)f_l - (k-1)f_k)T_b}$$
$$\times K_{\Pi}\left(t - \frac{(k+l-1)T_b}{2}, \tau - (k-l)T_b\right). \quad (24)$$

To remove the cross-terms, the cutoff Doppler-frequency of the TS function is set as $v_c \le |f_k - f_l|$. From (11), for this effect, the TS function parameter must be set such that

$$T_{sm} \ge \frac{3}{2|f_l - f_k|}. \quad (25)$$

However, for cross-terms between symbols of the same frequency, where $|f_k - f_l| = 0$, the TS function will not be able to remove them as they overlap with the autoterms.

Since FSK signal has two frequency components, the Doppler frequency is the difference between the two frequency components. For M-ary FSK signals, $|f_l - f_k|$ is set as the frequency deviation among the subcarrier frequencies. Any T_{sm} lower than the limit in (25) will not be able to remove the adjacent cross-terms, as the cutoff Doppler-frequency will include the cross-terms. For concentrated

autoterms, the low-pass filter should have a cutoff frequency that is as big as possible [11]. A high T_{sm} setting results in a small cutoff Doppler-frequency. This causes the autoterms to smear in time. For the best result, T_{sm} should be set just big enough to remove the cross-terms and not any bigger although the autoterms are concentrated at the Doppler axis to avoid smearing.

A balance choice of the values of T_g and T_{sm} will minimize the cross-terms while preserving the concentration of autoterms in the TFR [10]. For the FSK signal example, the TS function will remove the cross-term at symbols of $k = 2$ and $l = 3$ because the Doppler-frequency is nonzero when $f_k \neq f_l$. The rest of the cross-terms at symbols $k = 1$ and $l = 4$; $k = 2$ and $l = 3$; $k = 3$ and $l = 4$ and their reciprocal pairs can be removed by the lag-window.

4.3. Bilinear product of ASK signals

The time-lag representation for the ASK signal will also be represented based on the same binary sequence of "1101," for simplicity, and the modulation parameters defined in Section 2. For this signal, $f_k = f_0$ for $k = 1, 2, 4$ and $z(t) = 0$ for $k = 3$. For $k = 2$, the auto-term is

$$K_{z,2,2}\left(t - \frac{3T_b}{2}, \tau\right) = \exp(j2\pi f_0 \tau) K_{\Pi}\left(t - \frac{3T_b}{2}, \tau\right). \quad (26)$$

This function is centered at lag $\tau = 0$ and time $t = 3T_b/2$. The cross-term at $k = 1$ and $l = 4$ refers to the interaction between the 1st and the 4th symbols. Its bilinear product is expressed as

$$
\begin{aligned}
& K_{z,1,4}(t - 2T_b, \tau + 3T_b) \\
& = \exp\left(j2\pi f_0(\tau - 3T_b)\right) K_{\Pi}(t - 2T_b, \tau + 3T_b).
\end{aligned}
\quad (27)
$$

This cross-term is centered at $t = 2T_b$ and $\tau = -3T_b$. It is shown that there is a delayed lag-dependant component in this cross-term. The cross-term at $k = 2$ and $l = 3$, which refers to the interaction between the 2nd and the 3rd symbols, is expressed as

$$K_{z,2,3}(t - 2T_b, \tau + T_b) = 0 \cdot K_{\Pi}(t - 2T_b, \tau + T_b). \quad (28)$$

Since the 3rd symbol in this signal is $z_2(t) = 0$ (due to symbol "0"), then there is no cross-term here. The bilinear product representation of the ASK signal is shown in Figure 2.

The locations of autoterms and cross-terms are similar to the bilinear product of the FSK signal except that the cross-terms do not have Doppler-frequency components since there is only one subcarrier frequency present in ASK signal. The lag-window will be able to remove the components that lie away from the origin of the lag axis. By setting the lag-window width T_g as in (21), the autoterms are preserved while part of the cross-terms such as at $k = 1$ and $l = 2$ can be removed. Since the Doppler-frequency is zero and the lag-frequency is equal to the signal frequency, the remaining cross-terms do not introduce interference in the time-frequency representation.
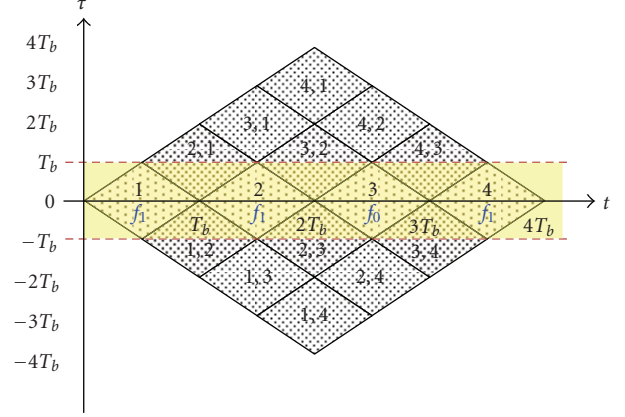

FIGURE 1: Bilinear product of FSK signal with lag-window. The bilinear products beyond the shaded area are removed.
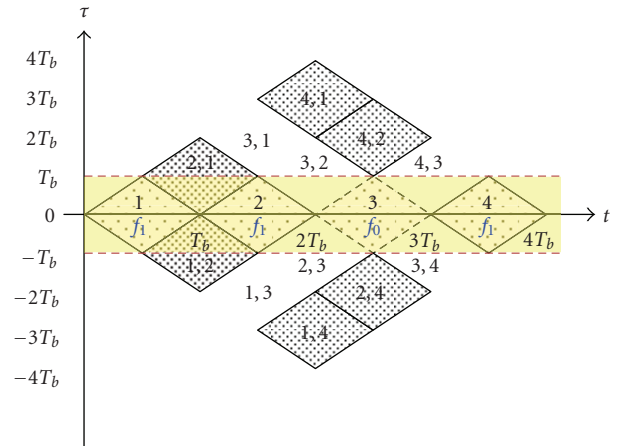

FIGURE 2: Bilinear product of ASK signal. The bilinear products beyond the shaded area are removed.

The use of the TS function in the SWWVD will not introduce any improvement in the TFR because all cross-terms have zero Doppler-frequency. Thus, the TS function property as a low-pass filter, in the Doppler-frequency domain, will pass all cross-terms. The TFD with only a lag-window, which is also known as window Wigner-Ville distribution (WWVD), is sufficient for ASK signals. In this paper, we use SWWVD on all the signals evaluated for uniformity. In this case, the TS function parameter T_{sm} is set to any small value so that it approaches an all-pass filter in the Doppler-frequency domain.

4.4. Kernel parameters

Based on (21) and (25), the limits of kernel parameters for various signals are summarized in Table 1. $T_{g,\max}$ is the largest lag-window width that can be set in (7) in order to obtain sufficient cross-terms reduction with minimal autoterms bias. $T_{sm,\min}$ is the smallest TS function parameter that can be set in (10) for the optimal representation. For
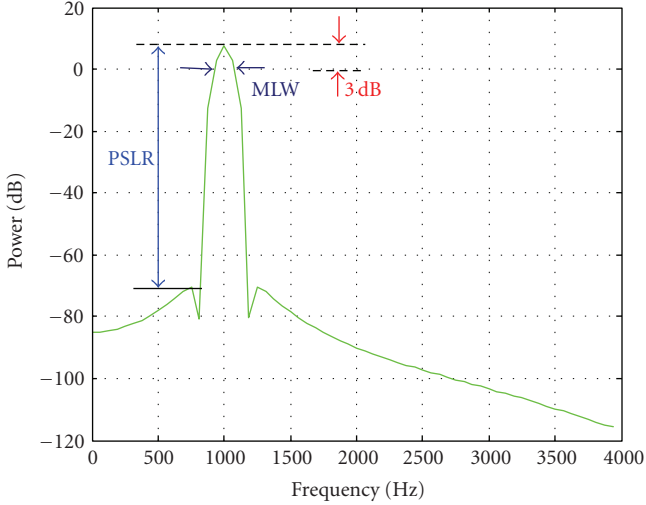
FIGURE 3: Performance measures used in the analysis, MLW and PSLR.

TABLE 1: Limit of kernel parameters. (Obtained mathematically from signal parameters.)

Signal	$T_{g,\max}$ (ms)	$T_{sm,\min}$ (ms)
FSK1	20.00	8.82
FSK2	13.33	8.82
FSK3	10.00	8.82
FSK4	8.000	8.82
ASK	10.00	—
8FSK	20.00	7.50
16FSK	20.00	7.50

(a) Time-frequency representation

(b) IF estimation

FIGURE 4: Time-frequency representation and the instantaneous frequency estimate from the TFR of ASK signal using $T_g = 10$ milliseconds, $T_{sm} = 10$ milliseconds.

ASK signal, T_{sm} can be set to any value, but preferably small so that the TS function approaches an all-pass filter.

To prove the limits in (21) and (25), we compare the performance of TFR with various kernels in terms main-lobe width (MLW), peak-to-side lobe ratio (PSLR), symbol-duration bias (SDB), and signal-to-cross-terms ratio (SCR). These performance measures are adopted and modified from [12], where they are used collectively to assess the performance of the TFDs in terms of its concentration, resolution, and interference minimization. In this paper, we compare the TFDs using each measure individually so that we will be able to see their effects independently.

MLW and PSLR are estimated from the power spectrum which is obtained from the frequency marginal of the TFR [13]. MLW is the width of the power spectrum, measured at 3 dB below the peak. Low MLW shows good frequency resolution as the peak is sharper and gives the ability to resolve closely spaced sinusoids. PSLR is the power ratio between the peak and the highest side-lobe, measured in dB. PSLR should be as high as possible to resolve signal of various magnitudes. The method to calculate MLW and PSLR is shown in Figure 3.

To calculate SDB, the estimated symbol-duration, which is obtained from the instantaneous frequency [13, 14] of the TFR, is compared with the actual symbol-duration of the transmitted signal:
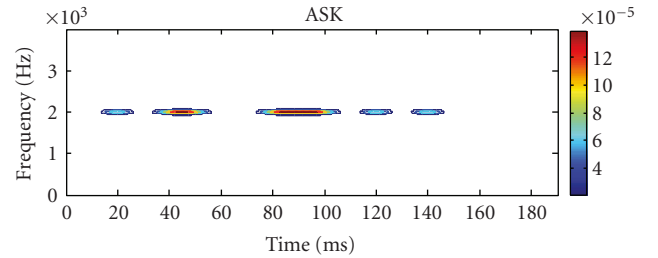
SDB

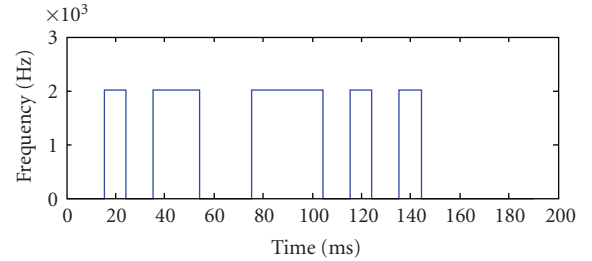$$= |\text{actual symbol-duration} - \text{estimated symbol-duration}|. \tag{29}$$

SDB shows the accuracy of the TFR in terms of time resolution of the digital communication signal. Previous TFD such as spectrogram suffers from bias in its representation. Its TFR fails to give the actual signal representation due to the tradeoff between its time and frequency resolution [2, 9]. An accurate time representation would give a biased frequency

resolution and vice versa in spectrogram. Low SDB shows that the TFR has good time resolution while low MLW shows that it has concentrated frequency resolution.

The volume of the TFR represents the energy of the signal. SCR is a ratio of autoterms power to cross-terms ratio in dB:

$$\text{SCR} = 10 \log \left(\frac{\text{signal power}}{\text{cross terms power}} \right). \tag{30}$$

High SCR shows high suppression of cross-terms in the TFR. In general, a good TFR should have low MLW and low SDB but high PSLR and high SCR.

The performance of TFD with various kernels is shown in Table 2. From Table 2, it is shown that for FSK2, the TFR is the optimal (low MLW, low SDB, high PSLR, and high SCR) when $T_g = 10$ milliseconds and $T_{sm} = 8.82$ milliseconds. Comparing with Table 1 which sets $T_{g,\max} = 13.33$ milliseconds and $T_{sm,\min} = 8.82$ milliseconds, smaller T_g gives better cross-terms suppressions which is seen in

TABLE 2: Performance comparison for various kernel parameters. (Main-lobe width (MLW) and symbol-duration bias (SDB) should be low but peak-to-side lobe ratio (PSLR) and signal-to-cross terms ratio (SCR) should be high. Area highlighted in blue shows the set of kernel parameters that give optimal representation of the signals. It is shown that there are a few sets of kernel parameters that can give optimal representation for each signal).

Kernel parameters	Performance measures	Signal						
		FSK1	FSK2	FSK3	FSK4	ASK	8FSK	16FSK
$T_g = 5$ ms, $T_{sm} = 10$ ms	MLW (Hz)	125.0	132.8	140.6	140.6	132.8	156.5	156.4
	SDB (ms)	0.406	0.483	0.625	0.406	0.125	0.250	0.188
	PSLR (dB)	5.652	5.467	6.278	6.325	6.460	5.271	5.297
	SCR (dB)	4.019	3.790	3.901	3.615	7.092	1.670	1.476
$T_g = 10$ ms, $T_{sm} = 5$ ms	MLW (Hz)	62.50	70.31	70.31	70.31	70.31	75.75	75.88
	SDB (ms)	2.000	7.400	1.400	6.500	0.125	0.250	0.313
	PSLR (dB)	5.455	5.427	5.303	5.646	6.129	5.234	5.293
	SCR (dB)	5.626	5.053	4.890	4.120	5.998	0.951	1.702
$T_g = 10$ ms, $T_{sm} = 8.82$ ms	MLW (Hz)	62.500	70.31	70.31	70.31	70.31	75.34	75.34
	SDB (ms)	0.469	0.514	0.688	0.531	0.125	0.125	0.188
	PSLR (dB)	5.718	6.672	5.422	5.663	6.129	5.318	5.308
	SCR (dB)	6.501	6.607	6.483	5.653	8.068	2.934	2.525
$T_g = 10$ ms, $T_{sm} = 10$ ms	MLW (Hz)	70.31	70.31	70.31	70.31	70.31	75.24	76.47
	SDB (ms)	0.406	0.483	0.781	0.531	0.125	0.250	0.188
	PSLR (dB)	5.721	6.726	5.405	5.567	6.129	5.249	5.291
	SCR (dB)	6.994	6.998	6.879	5.933	8.095	3.027	2.574
$T_g = 10$ ms, $T_{sm} = 12.5$ ms	MLW (Hz)	70.31	70.31	70.31	70.31	70.31	74.80	75.75
	SDB (ms)	0.406	0.483	1.200	3.300	0.250	0.250	0.188
	PSLR (dB)	5.702	6.704	5.378	5.801	6.130	5.231	5.282
	SCR (dB)	7.151	7.711	7.562	6.513	7.198	3.159	2.657
$T_g = 20$ ms, $T_{sm} = 5$ ms	MLW (Hz)	39.06	39.06	39.06	39.06	39.06	45.24	42.62
	SDB (ms)	12.50	7.100	4.900	1.800	6.100	0.250	0.531
	PSLR (dB)	5.824	5.334	5.472	5.329	6.830	5.459	5.247
	SCR (dB)	5.287	4.158	4.046	3.536	4.712	4.017	4.018
$T_g = 20$ ms, $T_{sm} = 7.5$ ms	MLW (Hz)	39.06	39.06	39.06	39.06	39.06	44.68	43.40
	SDB (ms)	0.281	9.240	9.600	3.200	2.700	0.250	0.594
	PSLR (dB)	5.833	5.326	5.835	5.957	5.599	5.418	5.316
	SCR (dB)	6.415	5.390	4.786	4.309	8.012	4.478	4.851
$T_g = 20$ ms, $T_{sm} = 8.82$ ms	MLW (Hz)	39.06	39.06	39.06	39.06	39.06	44.22	45.15
	SDB (ms)	0.375	2.360	9.800	5.400	9.900	0.125	0.188
	PSLR (dB)	5.843	5.238	5.614	5.252	6.026	5.234	5.240
	SCR (dB)	6.567	5.571	4.986	4.468	5.107	4.977	5.290

higher SCR but it suffers from increased MLW from the smearing of the autoterms. When $T_g = 10$ milliseconds but $T_{sm} < 8.82$ milliseconds, the MLW is similar but the SCR is smaller. This shows that the adjacent cross-terms are not reduced effectively, resulting in low SCR. However, the time resolution is good since the estimated symbol-duration is close to the actual (small SDB) as long as the parameter is not too small. Setting the T_{sm} to be too small will cause significant smearing of the autoterms in time direction, resulting in large SDB. As T_{sm} gets bigger, the SCR is higher as the adjacent cross-terms are removed. Although the SCR improves for large T_{sm}, the SDB gets worse as a result of smearing in the time representation. This is because the application of TS function in time domain is a convolution operation. Thus, there is a compromise between cross-terms suppression and time resolution.

At the optimal T_{sm}, when the lag-window is set such that $T_g < 10$ milliseconds, the SCR is higher because this window has shorter length and thus it can remove more cross-terms. However, it increases the MLW due to smearing of the autoterms in frequency direction, resulting in worsening the frequency resolution. Higher lag-window length at $T_g > 10$ milliseconds reduces the MLW and increases the autoterms concentration, at the expense of reduced SCR. The TFD has better frequency resolution but is unable to suppress cross-terms effectively, as more cross-terms are passed through the window. The presence of cross-terms in the TFR causes misinterpretation of the signal, resulting in higher SDB. Thus, there is a tradeoff between cross-terms suppression and frequency resolution.

Similar observations can be made on other signal models in this paper. The sets of kernel parameters that give the optimal TFR for each signal models are colored in Table 2.

A thorough performance analysis of FSK3 using various kernel parameters is shown graphically in Figure 5. By varying both kernel parameters, T_g and T_{sm}, an individual graph on PSLR, SCR, MLW, and SDB is derived. Analysis on these graphs shows that each performance measure is optimum at a different set of kernel parameters. The kernel parameters chosen must be able to give small MLW and SDB but large PSLR and SCR, at the same time. A balance must be made among these performance measures to achieve the optimal TFR. In our case, optimal kernel is set as a kernel with MLW $\leq 1/T_b$, SDB < 10%, PSLR > 5 dB, and SCR > 5 dB. PSLR and SCR are set to be more than 5 dB, so that the kernel will be able to give a reasonably good TFR in the presence of noise. For *M-ary* FSK signals, these criteria are relaxed in terms of PSLR and SCR, where they are set to be above 4 dB. Due to the presence of multi-subcarrier frequencies, the cross-terms in the TFR have more combination of the Doppler-frequencies.

From the analysis of Figure 5, it is shown that the kernel gives the best performance when $T_g = 10$ milliseconds and $T_{sm} = 8.82$ milliseconds. Further analysis of these parameters is performed by taking a slice at a-b, where $T_g = 10$ milliseconds and let T_{sm} vary. The performance of FSK3 under various T_{sm} values is then plotted in Figure 6(a). It is shown that the optimal T_{sm} falls between 8 milliseconds and 12 milliseconds. Next, the slice at c-d is observed for $T_{sm} = 8.82$ milliseconds and let T_g vary, all the performance measures are shown in Figure 6(b). It is shown that the optimal T_g falls between 9 milliseconds and 11 milliseconds. These findings are comparable to the results shown in Table 2 and consistent for all the signals in this paper. It is observed that the TFR is optimal when T_g and T_{sm} are approximate to the limits given in (21) and (25).

From the findings, we can see that the limits at (21) and (25) can be used as the guideline for optimal TFR. For some cases, a small variation from the parameters limit gives better overall performance than at the limit itself. This is because the optimality is seen as an overall performance and each performance measure changes differently with the change of kernel parameters. There are few sets of kernel parameters that can give the optimal TFR as there is no one specific kernel parameters that can give the optimal performance for every performance measure used in this paper, simultaneously. Thus, we can conclude that for the best TFR, the lag-window width and the TS function parameter should be set as close as possible to the limit $T_{g,\max}$ and $T_{sm,\min}$. However, for an adaptive optimal kernel system which will function without prior knowledge of the signal, we set the lag-window width as $T_{g,\max}$ and the TS function parameter as $T_{sm,\min}$ so that a small bias in the signal parameters estimate (this is likely to happen due to the smoothing method (LOWESS) which will be discussed in Section 5) will not affect the overall performance of the TFR. For ASK signal, any value of T_{sm} set as the TS function does not affect the performance. Based on previous discussion, T_{sm} should be small so that the TS function approaches a Dirac delta function. In this case, we set $T_{sm} = T_g = 10$ milliseconds.

4.5. Computation complexity

Assuming perfect knowledge of the signal, the number of computation required to implement the optimal SWWVD in terms of number of multiplication is given as [15]

(1) bilinear product requires $N_\tau N$ multiplications;

(2) product between bilinear product and the lag window requires $N_\tau N$ multiplications;

(3) convolution with the setup time-smooth function requires $N_{sm}N_\tau N$ multiplications;

(4) Fourier transform of the time-lag representation requires $0.5 N_{2\tau} N \log_2 N_{2\tau}$ multiplications,

where N is the signal length, N_τ is the window length, N_{sm} is the length of smoothing function, and $N_{2\tau}$ is the length that is multiple of 2 and is greater or equal to N_τ. Thus, the total of multiplication required to compute SWWVD is $(2N_\tau + N_{sm}N_\tau + 0.5N_{2\tau}\log_2 N_{2\tau})N$.

4.6. Guideline to determine kernel parameters

Based on the analysis above, a guideline to determine the separable kernel parameters for SWWVD is given as follows.

(i) For multifrequency signal,

$$|T_g| = T_b, \qquad T_{sm} = \left[\frac{3}{2|f_l - f_k|}\right]. \tag{31}$$

(ii) For single-frequency signal,

$$|T_g| = T_b, \qquad T_{sm} = T_g. \tag{32}$$

5. ADAPTIVE OPTIMAL KERNEL

Based on the analysis made in Section 4, we design an adaptive optimal kernel SWWVD that is capable of giving an optimal TFR without having prior knowledge of the signal. In this system, first, the signal parameters such as symbol-duration and subcarrier frequencies will be estimated from the input signal. These parameters will be used to design the optimal kernel for this signal. Symbol-duration is determined from the autocorrelation function while the subcarrier frequencies are determined from the spectrum of the signal.

Since the kernel parameters can be set from the signal parameters such as symbol-duration and subcarrier frequencies, these parameters must be estimated from the input signal before the TFR is calculated. Symbol-duration of a random process can be estimated from the autocorrelation of the signal and the subcarrier frequencies can be obtained from the energy spectrum.

5.1. System design

Figure 7 shows the system design of the adaptive SWWVD. For any unknown signal, first the bilinear product will be calculated. From the bilinear product, the autocorrelation

(a) Optimal kernel gives maximum PSLR

(b) Optimal kernel gives maximum SCR

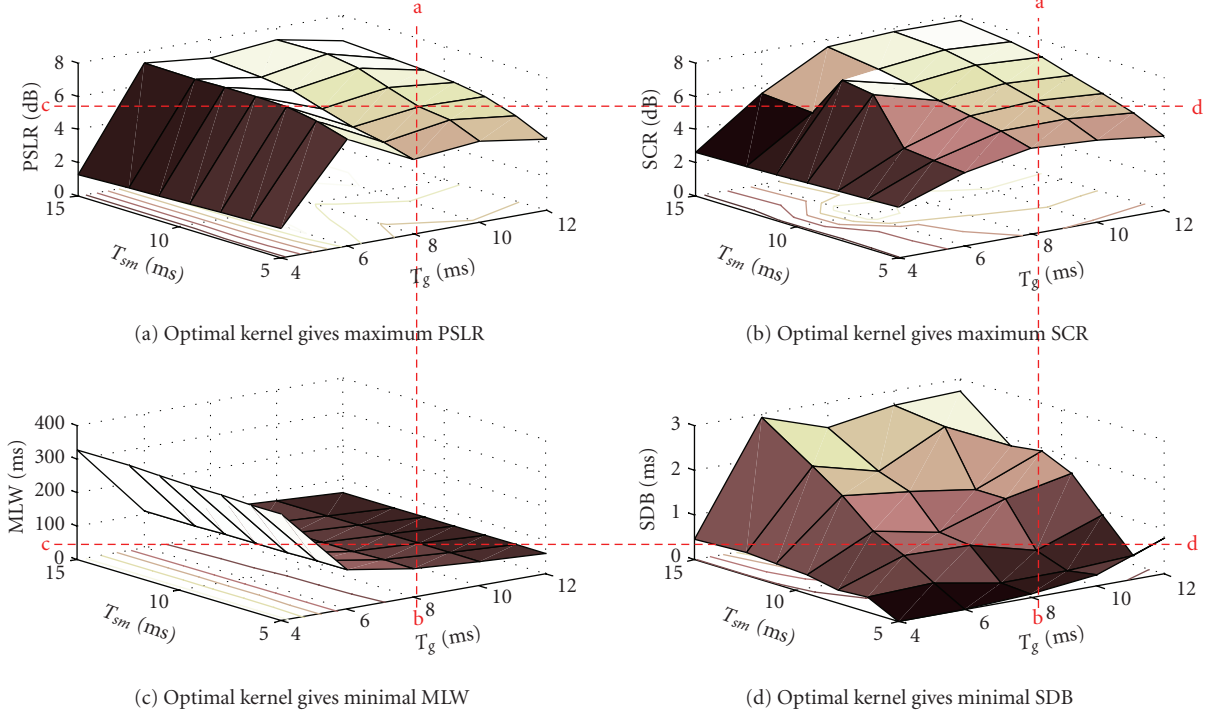(c) Optimal kernel gives minimal MLW

(d) Optimal kernel gives minimal SDB

FIGURE 5: Performance of the TFR of FSK3 using various kernel parameters. (The optimal kernel is chosen from the kernel parameters that give small MLW and small SDB but large PSLR and large SCR, simultaneously.)

function will be obtained. Based on the autocorrelation function, the symbol-duration and subcarrier frequencies will be estimated. These parameters are then used to design the optimal kernel for this signal.

5.2. Bilinear product

For any unknown signal that is input into the system, we will first calculate the bilinear product $K_z(t, \tau)$ given in (12)–(17). The autocorrelation function of the signal can then be calculated from the bilinear product.

5.3. Autocorrelation function

For a digital modulation signal with a pseudorandom sequence, the autocorrelation function is given as [16]

$$R_z(\tau) = \int K_z(t, \tau)dt = \left[1 - \frac{|\tau|}{T_b} \right] \sum_{k=0}^{N-1} \exp(j2\pi f_k \tau), \quad (33)$$

where T_b is the symbol-duration or the period of the signal, N is the number of symbol in the binary sequence, and f_k is the subcarrier frequency at kth symbol. The autocorrelation function, $R_z(t, \tau)$, provides a measure on how closely the signal matches a copy of itself as the copy is shifted τ units in time [16]. The autocorrelation function of a periodic signal is also periodic at the period similar to the signal.

5.4. Smoothing

In order to determine the symbol-duration, first we need to find the envelope of the autocorrelation function. The autocorrelation function satisfies the symmetry condition indicated in the following:

$$R_z^*(\tau) = R_z(-\tau). \quad (34)$$

Consequently, we need to consider the envelope for positive lag values only. It is obtained by multiplying the autocorrelation function with its conjugate:

$$E_{Rz}(\tau) = R_z(\tau) \cdot R_z^*(\tau) = \left[1 - \frac{|\tau|}{T_b} \right]^2. \quad (35)$$

However, due to the nonrandomness of the sequence and the limitation of signal length in the signal that we evaluate, the envelope that is obtained has some out-of-correlation terms. The envelope has to be smoothed before it is used for estimating the symbol-duration. In this paper, we use the locally weighted regression (LOWESS), which is discussed in detail in [17], as the smoothing function. This method is an extension of the weighted least squares (WLSs) to locally smoothing the scatterplots. The cost function of the weighted linear regression is given as [18]

$$J_z(\tau) = ||\mathbf{E}_{Rz} - \boldsymbol{\beta}\boldsymbol{\tau}||_W^2 \triangleq (\mathbf{E}_{Rz} - \boldsymbol{\beta}\boldsymbol{\tau})^* \mathbf{W}(\mathbf{E}_{Rz} - \boldsymbol{\beta}\boldsymbol{\tau}), \quad (36)$$

where $||\cdot||^2$ denotes the squares Euclidean norm, $\mathbf{W}$ is any Hermitian positive-definite weight function, $\boldsymbol{\beta}$ is an unknown $N \times n$ matrix, $\mathbf{E}_{Rz}$ is an $N \times 1$ vector of the envelope,

(a1) MLW and SDB of FSK3 under various T_{sm} at fixed $T_g = 10$ ms

(a2) PSLR and SCR of FSK3 under various T_{sm} at fixed $T_g = 10$ ms

(a)

(b1) MLW and SDB of FSK3 under various T_g at fixed $T_{sm} = 8.82$ ms

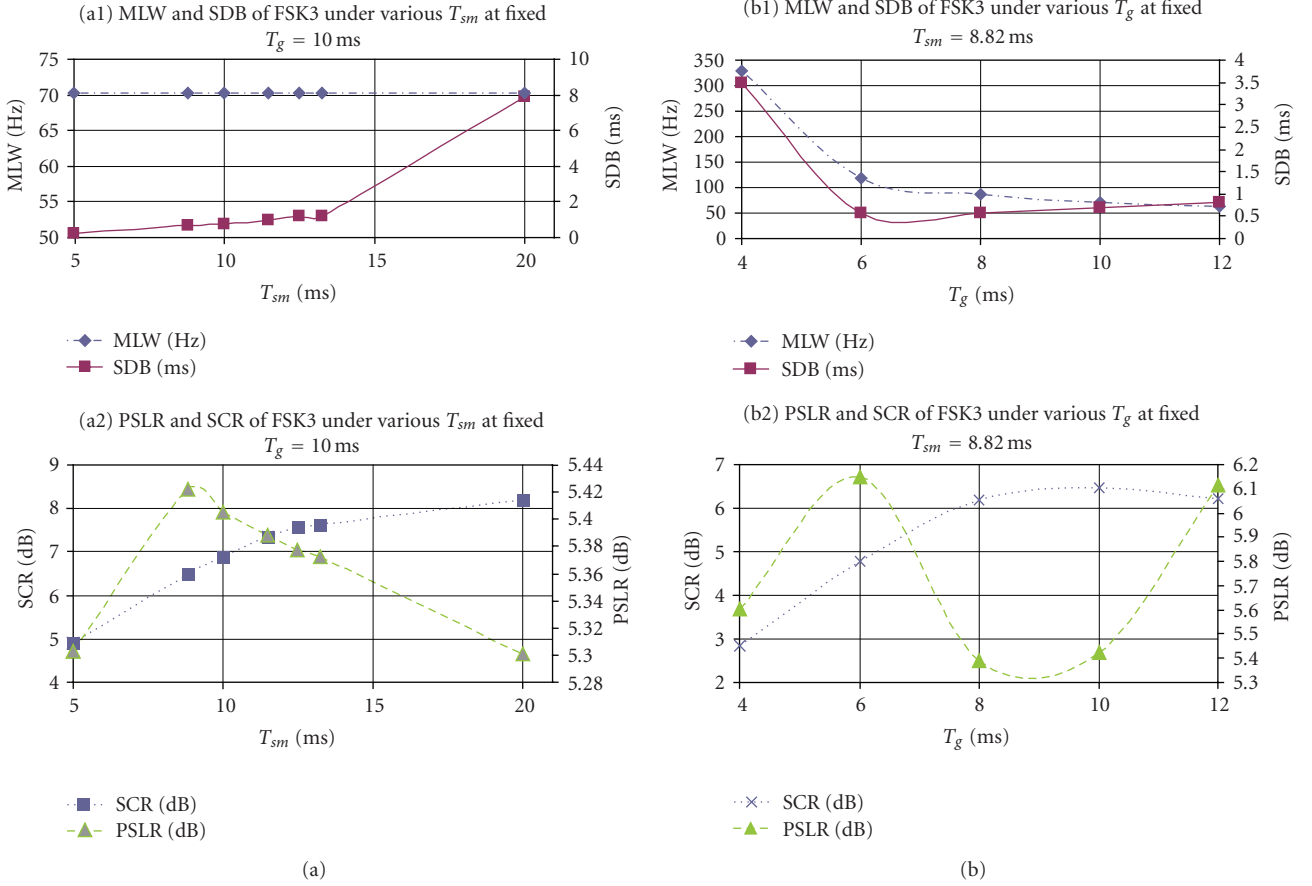(b2) PSLR and SCR of FSK3 under various T_g at fixed $T_{sm} = 8.82$ ms

(b)

FIGURE 6: Performance of the TFR of FSK3 using various kernel parameters. (Optimal kernel is the kernel with a small MLW and small SDB but big SCR and big PSLR).
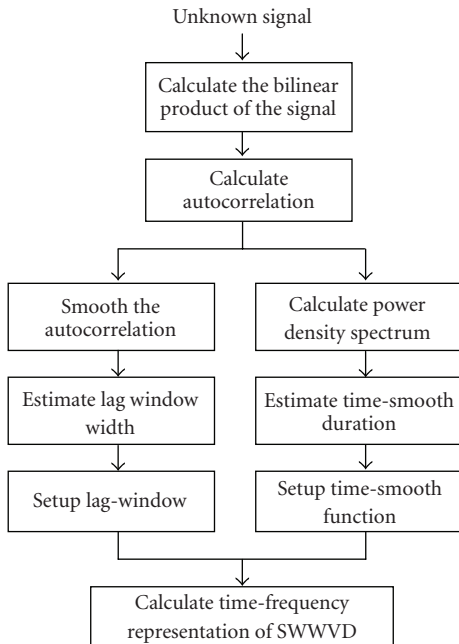

FIGURE 7: System design of adaptive SWWVD.

and $\boldsymbol{\tau}$ is an $n \times 1$ vector. For normal linear regression, $\mathbf{W}$ matrix is equals to 1. In most applications, we would like to find the optimum $\widehat{\boldsymbol{\beta}}$ which minimizes the cost function

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{\tau} \mathbf{W} \boldsymbol{\tau}^{\mathbf{T}})^{-1} \boldsymbol{\tau} \mathbf{W} \mathbf{E}_{R\mathbf{z}}, \quad \mathbf{W} = \text{diag}\{\omega_1, \dots, \omega_n\}. \quad (37)$$

In WLS, $\widehat{\boldsymbol{\beta}}$ is estimated for a given block of observed data. However in LOWESS, $\widehat{\boldsymbol{\beta}}$ is calculated at every τ, where three main steps are carried out (1) determine the weights of all time instances k, relative to τ; (2) estimate $\widehat{\beta}_0(\tau)$ and $\widehat{\beta}_1(\tau)$, and (3) calculate the smoothed curve. Given an envelope of the autocorrelation function, E_{Rz} for $0 \leq \tau \leq T$, first, the weights are calculated. The weights can be seen as a window function, which is given as

$$w_k(\tau) = \left(1 - \left|\frac{k - \tau}{T_r}\right|^3\right)^3, \quad \text{for } |k - \tau| < T_r,$$
$$= 0, \quad \text{for } |k - \tau| \geq T_r. \quad (38)$$

At one particular $\tau = \tau_1$, the weight $w_k(\tau_1)$ is centered at $k = \tau_1$ and the weights are calculated for all time instances,

k, $0 \le k \le T_r$ using (38). From simulation, it is shown that for the type of signals in this paper, it is best to set as $T_r = 9.375$ milliseconds. Once the weights are calculated, they are used as the weight function in the WLS at $\tau = \tau_1$. The estimated envelope is given as

$$\hat{E}_{Rz}(k) = \beta_0(\tau) + \beta_1(\tau)k. \tag{39}$$

The error between actual and estimated envelope is given as

$$e(k) = E_{Rz}(k) - \hat{E}_{Rz}(k) = E_{Rz}(k) - \beta_0(\tau) - \beta_1(\tau)k, \tag{40}$$

for $0 \le k \le T$. The cost function of the localized WLS with weight $w_k(\tau)$ for $(k, E_{Rz}(k))$, $J_z(\tau)$ is given as

$$
\begin{aligned}
J_z(\tau) &= E\left[\int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)e^2(k)dk\right] \\
&= E\left[\int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)(E_{Rz}(k) - \beta_0(\tau) - \beta_1(\tau)k)^2 dk\right],
\end{aligned}
$$

$$\tag{41}$$

$$
\begin{pmatrix} \hat{\beta}_0(\tau) \\ \hat{\beta}_1(\tau) \end{pmatrix} =
\begin{pmatrix}
\int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)dk & \int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)k\,dk \\
\int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)k\,dk & \int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)k^2 dk
\end{pmatrix}^{-1}
$$
$$
\times \begin{pmatrix}
\int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)E_{Rz}(k)dk \\
\int_{k=\tau-T_r}^{k=\tau+T_r} w_k(\tau)E_{Rz}(k)k\,dk
\end{pmatrix}.
$$

$$\tag{42}$$

Thus $\hat{\beta}_0(\tau)$ and $\hat{\beta}_1(\tau)$ are the values of $\beta_0(\tau)$ and $\beta_1(\tau)$ that minimize (41). The smoothed point at τ_1 using locally WLS is $(\tau_1, \hat{E}_{Rz}(\tau_1))$, where $\hat{E}_{Rz}(\tau_1)$ is the fitted value of the regression at $\tau = \tau_1$

$$\hat{E}_{Rz}(\tau) = \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau)\tau. \tag{43}$$

The window function is then slid and centered at next τ, noted as τ_2 and then the weights $w_k(\tau_2)$ for all time instant, k, $\tau_2 - T_r \le k \le \tau_2 + T_r$ are calculated. Next, $\hat{\beta}_0(\tau_2)$ and $\hat{\beta}_1(\tau_2)$ are estimated from the WLS and then, the estimated $\hat{E}_{Rz}(\tau_2)$ is calculated. These steps are repeated for all τ, $0 \le \tau \le T$. The autocorrelation function, envelope of the autocorrelation function and envelope after LOWESS smoothing, is shown in Figure 8.

The first crossing of the smoothed envelope is the estimated symbol-duration, $\hat{T}_b$, which is used to set the lag-window in (7) by equating $T_g = \hat{T}_b$. This lag-window is then used in generating the TFR of the signal.

(a) Autocorrelation function of FSK signal

(b) Envelope of the autocorrelation function

First crossing = symbol-duration
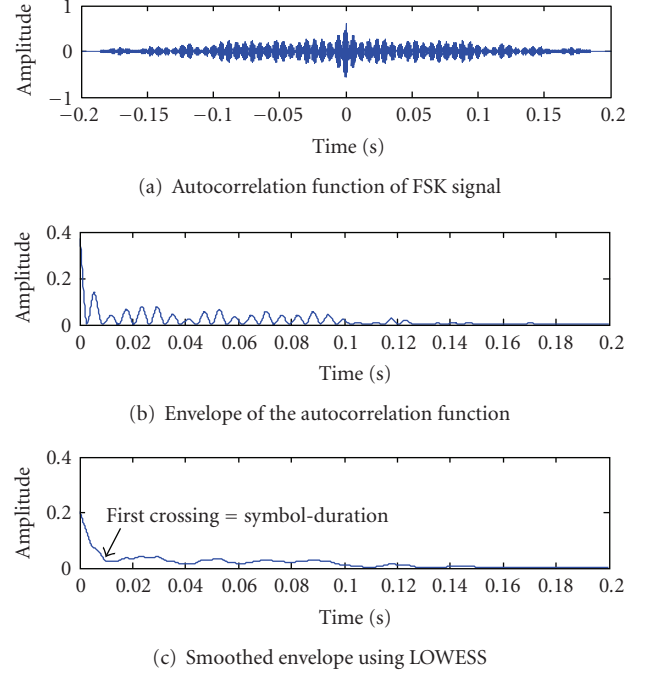
(c) Smoothed envelope using LOWESS

FIGURE 8: The autocorrelation function of FSK signal and the smoothed curve for symbol-duration estimate.

TABLE 3: Optimum and estimated kernel parameters.

Signal	Optimum kernel		Estimated kernel	
	$T_{g,opt}$ (ms)	$T_{gsm,pt}$ (ms)	$T_{g,estt}$ (ms)	$T_{gsm,est}$ (ms)
FSK1	20.00	8.820	18.70	8.800
FSK2	13.33	8.820	13.10	8.800
FSK3	10.00	8.820	8.400	8.800
FSK4	8.000	8.820	8.300	8.800
ASK	10.00	10.00	9.400	10.00
8FSK	20.00	7.500	18.70	8.600
16FSK	20.00	7.500	18.70	8.600

5.5. Power density spectrum

According to the Wiener-Khinchine theorem, the power density spectrum is the Fourier transform of the autocorrelation function [16]. The power spectrum obtained from the estimated autocorrelation function in (33) is

$$S_z(f) = \underset{\tau \to f}{FT}\{R_z(\tau)\} = \int_{-\infty}^{\infty} R_z(\tau)e^{-j2\pi f\tau}d\tau. \tag{44}$$

From the power density spectrum, the frequency content of the signal can be estimated. The smallest difference in the subcarrier frequencies, Δf, is used as the denominator to T_{sm} at

$$T_{sm} = \left[\frac{3}{2}|\Delta f|\right]. \tag{45}$$

The value of T_{sm} is then replaced into (10). For ASK signals, where there is only one subcarrier frequency, T_{sm} is

TABLE 4: Performance comparison between actual kernel and estimated kernel.

Signal	Performance measures	Without noise		SNR = 12 dB	
		Optimal kernel	Adaptive kernel	Optimal kernel	Adaptive kernel
FSK1	MLW (Hz)	39.06	39.06	39.06	39.06
	SDB (ms)	0.375	0.250	0.288	0.283
	PSLR (dB)	5.843	7.093	5.842	7.054
	SCR (dB)	6.507	6.570	6.565	6.569
FSK2	MLW (Hz)	54.69	54.69	54.68	54.69
	SDB (ms)	0.514	0.514	0.536	0.536
	PSLR (dB)	5.560	5.538	5.557	5.541
	SCR (dB)	6.370	6.319	6.370	6.316
FSK3	MLW (Hz)	70.31	85.94	70.31	85.94
	SDB (ms)	0.688	0.625	0.693	0.614
	PSLR (dB)	5.422	5.574	5.567	5.573
	SCR (dB)	6.483	6.428	6.370	6.425
FSK4	MLW (Hz)	85.94	85.44	85.43	83.59
	SDB (ms)	0.469	0.500	0.463	0.500
	PSLR (dB)	5.587	5.285	5.585	5.415
	SCR (dB)	5.507	5.570	5.506	5.580
ASK	MLW (Hz)	70.31	70.31	70.94	70.31
	SDB (ms)	0.025	0.025	0.050	0.050
	PSLR (dB)	6.129	5.543	6.128	5.544
	SCR (dB)	8.095	8.543	8.091	8.536
8FSK	MLW (Hz)	44.63	44.81	45.22	45.08
	SDB (ms)	0.025	0.125	0.190	0.143
	PSLR (dB)	5.418	5.590	5.407	5.603
	SCR (dB)	4.478	4.785	4.407	4.730
16FSK	MLW (Hz)	43.20	43.73	43.20	43.73
	SDB (ms)	0.532	0.188	0.566	0.443
	PSLR (dB)	5.316	5.274	5.327	5.343
	SCR (dB)	4.853	4.778	4.726	4.616

set as 10 milliseconds. This TS function is then used in the SWWVD.

5.6. Computation complexity

To implement the adaptive SWWVD, the number of computations in terms of multiplication that is required is [15]

(1) bilinear product and autocorrelation function require $N_\tau N$ multiplications;

(2) smoothing the autocorrelation (LOWESS) requires N_τ times the computation as follows:

    (a) weights function requires $12 N_r$ multiplications;
    (b) inverse matrix of (43) requires $8(6N_r + 6N_r^2)$ multiplications;

(3) power density spectrum (Fourier transform of autocorrelation function) requires $0.5 N_\tau N_{2\tau} \log_2 N_{2\tau}$ multiplications;

(4) product between bilinear product and the setup lag window requires $N_\tau N$ multiplications;

(5) convolution with the setup time-smooth function requires $N_{sm} N_\tau N$ multiplications;

(6) Fourier transform of the time-lag representation requires $0.5 N_{2\tau} N \log_2 N_{2\tau}$ multiplications;

where N_r is the length of window function used in the weight function. The total number of multiplication required to implement the adaptive SWWVD is $(2N_\tau + N_{sm}N_\tau + 0.5N_{2\tau}\log_2 N_{2\tau})N + N_\tau[12N_r(4N_r+5) + 0.5N_{2\tau}\log_2 N_{2\tau}]$. The adaptive SWWVD requires an additional multiplication of $N_\tau[12N_r(4N_r+5) + N_{2\tau}/2(\log_2 N_{2\tau})]$ than the computation of the optimal SWWVD. However, in noncooperative environment, this method is an advantage as it requires no prior knowledge of the signal and yet able to give the optimal representation of the signal.

6. RESULTS

Assuming perfect prior knowledge of the signal, the optimal kernel parameters are determined from (31) and (32) in Section 4.4 and summarized in Table 3. Then they are compared with the adaptive kernels which are obtained from the estimated signal parameters as discussed in Section 5. The signals are assumed to be under ideal conditions with no interference due to noise. It is observed that the estimated kernel parameters are close to the optimum parameters. In general, the estimated lag-window width is smaller than the optimum while the estimated time-smooth duration approaches the optimum. A small variation in the estimated value is due to the smoothing method (LOWESS) used in the estimation of symbol-duration and subcarrier frequencies in the adaptation process. However, it does not affect the TFR significantly, as shown in Table 4.

To benchmark the adaptive kernel SWWVD against the optimal kernel SWWVD, we compare the performance for every signal at ideal case (without noise), and at SNR of 12 dB. This is because most of the signal classification techniques evaluate signals at SNR between 10 dB and 20 dB [19]. A Monte Carlo simulation based on 100 realizations was conducted for each signal to evaluate the robustness of the TFR in the presence of noise. At every realization, additive white Gaussian noise is added to the signal and the performance of the TFR is noted. The performance measures are then averaged and tabulated in Table 4. They are compared in terms of MLW, PSLR, SDB, and SCR. On the average, it is observed that the performance of this adaptive system is comparable to the optimal system.

Next, for further analysis, the performance between optimal SWWVD and adaptive SWWVD is compared for FSK1 with SNR from −2 dB to 12 dB. This comparison is shown graphically in Figure 9. At SNR > 0 dB, both optimal SWWVD and adaptive SWWVD perform at comparable results in terms of MLW, SDB, and SCR. The adaptive system outperforms the optimal system in terms of PSLR. It is shown that the performance difference between the optimal kernel and adaptive kernel is insignificant and negligible.

Both the optimal system and the adaptive system start to fail when SNR < 0 dB. MLW and SDB start to get bigger while the PSLR and SCR start to get smaller than the assigned limits. The TFR at SNR < 0 dB will not be able to give the correct representation of the signal and thus causes misinterpretation of the signal. Similar TFR performance of the optimal SWWVD and the adaptive SWWVD is observed for the other signals presented in this paper.

Performance comparability can also be seen in Figure 10, where the TFR of FSK3 is obtained using optimal SWWVD and adaptive SWWVD. The adaptive SWWVD shows a cleaner TFR which indicate higher PSLR and SCR. All in all, we can conclude that the performance of the adaptive SWWVD is comparable to the optimal SWWVD even in noisy conditions.

Next, the computation complexity is compared. In this paper, the evaluated signal has the signal length N of 1024 sample points and the lag-window length N_τ is set as 512 sample points. $N_{2\tau}$ is also equal to 512 sample points,
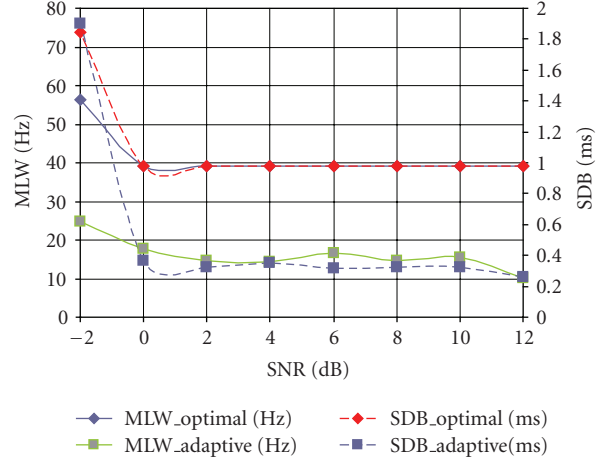
(a) MLW and SDB comparison between optimal kernel and adaptive kernel for FSK1
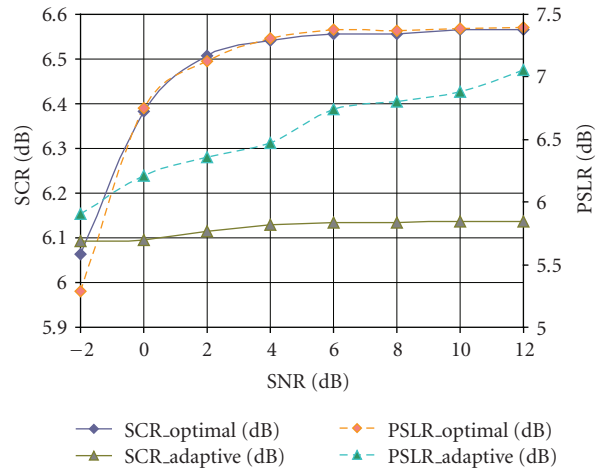
(b) PSLR and SCR comparison between optimal kernel and adaptive kernel for FSK1

FIGURE 9: Performance comparison between optimal kernel and adaptive kernel for FSK1 in the presence of various SNR levels.

for this case. At worse case condition, the time-smooth function N_{sm} is of 80 sample points (ASK signal). The weight function N_r is set to be 75 sample points. From the computation calculation in Section 4.5, the SWWVD requires 4.535×10^7 multiplications to implement. On the other hand, from Section 5.6, the adaptive SWWVD requires 1.871×10^8 multiplications, which is approximately 4 times more computations as compared the optimal SWWVD. This is the worst case condition, assuming that the computation of LOWESS is not carried out using the optimized methods. An improved method which requires less computation has been developed [17].

Although there is a significant additional multiplication in the adaptive SWWVD, the improvement in the time-frequency representation of signals in noncooperative environment enables more efficient signal analysis. On top of that, with the current advancement in the computer processors, the adaptive SWWVD can be implemented without causing much delays in the processing time by using
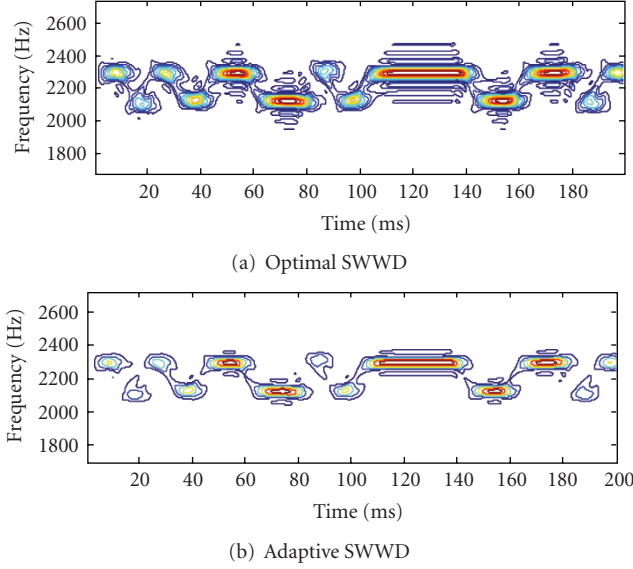
(a) Optimal SWWD

(b) Adaptive SWWD

FIGURE 10: The TFR using optimal SWWVD and using adaptive SWWVD of FSK3.

TABLE 5: Instantaneous autocorrelation function of the box function.

Lag shift, τ	Time instant when $K_\Pi(t, \tau) = 1$
T_b	$t = T_b/2$
$T_b/2$	$T_b/4 \leq t \leq 3T_b/4$
0	$0 \leq t \leq T_b$
$-T_b/2$	$T_b/4 \leq t \leq 3T_b/4$
$-T_b$	$t = T_b/2$

multiple processors, parallel processing, and so on [20]. Other types of faster smoothing function in [18, 21, 22] and the references in [22] can also be implemented to replace LOWESS.

7. CONCLUSIONS

Signals with different signal parameters, that is, subcarrier frequencies and symbol-duration need kernel with different parameters for optimal TFR. There is no single kernel that can be used optimally for all signals, even for signal of the same modulation type but with different signal parameters. The optimal kernel parameters can be determined mathematically from the signals; T_{sm} length is the inverse of the frequency deviation and the lag-window width T_g is symbol-duration of the signal. By determining the kernel parameters, the optimum distribution can be designed.

The performances of the system with adaptive kernel and optimal kernel are comparable in terms of MLW, PSLR, SDB, and SCR. This shows that the adaptive system designed can be used to obtain optimal TFR automatically for signal in noncooperative environment at the cost of some additional multiplications.

APPENDICES

The bilinear product, which can be seen as the instantaneous correlation function, is given in (4). For a short-duration complex exponential signal,

$$z(t) = A \exp(j2\pi f_1 t)\Pi(t), \tag{A.1}$$

where $\Pi(t)$ is the box function given in (2), the instantaneous autocorrelation function (IAF) is given as

$$K_z(t, \tau) = |A|^2 \exp(j2\pi f \tau)\Pi\left(t + \frac{\tau}{2}\right)\Pi\left(t - \frac{\tau}{2}\right)$$
$$= |A|^2 \exp(j2\pi f \tau)K_\Pi(t, \tau), \tag{A.2}$$

where T_b is the signal duration, A is the amplitude, and $K_\Pi(t, \tau)$ is the IAF of the box-function given as

$$K_\Pi(t, \tau) = \Pi\left(t + \frac{\tau}{2}\right)\Pi\left(t - \frac{\tau}{2}\right). \tag{A.3}$$

IAF provides a measure on how closely the signal matches a copy of itself as the copy is shifted τ units in time [16]. For the IAF to have the maximum value, the signals must be identical to each other at that instant. For a box function, this condition is true when they occur at the same time instant when the box function overlaps a copy of itself, $\tau = 0$:

$$t + \frac{\tau}{2} = t - \frac{\tau}{2}, \quad \tau = 0. \tag{A.4}$$

The IAF has the same duration as the duration of the box function:

$$K_\Pi(t, \tau)|_{\tau=0} = 1, \quad \text{when } 0 \leq t \leq T_b,$$
$$= 0, \quad \text{elsewhere.} \tag{A.5}$$

From (A.3), when the box function is shifted by $\tau = T_b/2$, the box function is shifted such that it lags by $T_b/4$ while its copy lags by $T_b/4$. Thus, they overlap each other when $T_b/2 \leq t \leq 3T_b/2$. As the shift gets bigger, the overlap between the box function and its copy gets smaller. The same scenario is seen when the box function is shifted by $\tau \leq 0$. The summary of the IAF of the box function is given in Table 5. The IAF has zero value elsewhere. From Table 5, it is observed that the IAF has a rhombic shape of the same duration as the box function $0 \leq t \leq T_b$ and is centered at the origin of lag axis, where it has values for $-T_b \leq \tau \leq T_b$.

An arbitrary digital communication signal can be seen as a sum of N short-duration complex exponential signal given in (1). The bilinear product of this signal can be seen as a combination of N-delayed IAF, which is located along the time axis and centered at the origin of lag axis.

However, the interaction between symbols of different frequency introduces additional bilinear product known as cross-terms, which is located away from the time axis at $|\tau| \geq 0$ [9]. These additional terms are not a representation of the actual signal and thus, if they are not removed, they will cause misinterpretation in the TFR. To illustrate this, we evaluate

an FSK signal of symbol sequence "1101" with the amplitude, $A = 1$. This signal is given as

$$z(t) = \sum_{k=1}^{4} A_k \exp\left(j2\pi f_k(t - (k-1)T_b)\right)\Pi(t - (k-1)T_b)$$

$$= \exp(j2\pi f_1 t)\Pi(t) + \exp\left(j2\pi f_1(t - T_b)\right)\Pi(t - T_b)$$

$$+ \exp\left(j2\pi f_0(t - 2T_b)\right)\Pi(t - 2T_b)$$

$$+ \exp\left(j2\pi f_1(t - 3T_b)\right)\Pi(t - 3T_b).$$

$$\text{(A.6)}$$

From (5), the bilinear product is given in (A.5)

$$K_z(t, \tau) = z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right)$$

$$= \sum_{k=1}^{4} A_k \exp\left(j2\pi f_k\left(t + \frac{\tau}{2} - (k-1)T_b\right)\right)$$

$$\times \Pi\left(t + \frac{\tau}{2} - (k-1)T_b\right)$$

$$\times \sum_{l=1}^{4} A_l^* \exp\left(-j2\pi f_l\left(t - \frac{\tau}{2} - (l-1)T_b\right)\right)$$

$$\times \Pi^*\left(t - \frac{\tau}{2} - (l-1)T_b\right),$$

$$K_z(t, \tau) = \sum_{k=1}^{4}\sum_{l=1}^{4} \exp(j2\pi(f_k - f_l)t) \exp\left(j2\pi(f_k + f_l)\frac{\tau}{2}\right)$$

$$\times \exp\left(j2\pi((l-1)f_l - (k-1)f_k)T_b\right)$$

$$\times \Pi\left(t + \frac{\tau}{2} - (k-1)T_b\right)\Pi^*\left(t - \frac{\tau}{2} - (l-1)T_b\right),$$

$$K_z(t, \tau) = \sum_{k=1}^{4}\sum_{l=1}^{4} \exp\left(j2\pi(f_k - f_l)t\right) \exp\left(j2\pi(f_k + f_l)\frac{\tau}{2}\right)$$

$$\times \exp\left(j2\pi((l-1)f_l - (k-1)f_k)T_b\right)$$

$$\times K_\Pi\left(t - \frac{(k+l-1)T_b}{2}, \tau - (k-l)T_b\right),$$

$$K_z(t, \tau) = \sum_{k=1}^{4} |A_k|^2 \exp(j2\pi f_k \tau) K_\Pi\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right)$$

$$+ \sum_{\substack{k=1,\, l=1 \\ k \neq l}}^{4} A_k A_l^* \cdot \exp\left(j2\pi((k-1)f_k - (l-1)f_l)T_b\right)$$

$$\times \exp\left(j2\pi\left(\frac{(f_k + f_l)}{2}\right)\tau\right) \exp\left(j2\pi(f_l - f_k)t\right)$$

$$\times K_\Pi\left(t - \left(\frac{(k+l-1)T_b}{2}\right), \tau - (k-l)T_b\right)$$

$$= K_{z,\text{auto}}(t, \tau) + K_{z,\text{cross}}(t, \tau).$$

$$\text{(A.7)}$$

This bilinear product of the box function has a rhombic shape and centered at $t = (k + l)T_b$ and $\tau = (k - l)T_b$. The bilinear product of this signal has both autoterms and cross-terms.

A. AUTOTERMS

The autoterms are bilinear products of signals with the same symbol. For example, the IAF of symbol sequence 1 is given as follows:

$$K_{z,1,1}(t, \tau) = e^{j2\pi f_1(t+\tau/2)}\Pi\left(t + \frac{\tau}{2}\right) \cdot e^{-j2\pi f_1(t-\tau/2)}\Pi\left(t - \frac{\tau}{2}\right)$$

$$= e^{j2\pi f_1 \tau} K_\Pi\left(t - \frac{T_b}{2}, \tau\right).$$

$$\text{(A.8)}$$

For an arbitrary kth symbol, the IAF is

$$K_{z,k,k}\left(t - \frac{(2k-1)T_b}{2}, \tau\right)$$

$$= e^{j2\pi f_k(t+\tau/2-(k-1)T_b)}\Pi\left(t + \frac{\tau}{2} - (k-1)T_b\right)$$

$$\times e^{-j2\pi f_k(t-\tau/2-(k-1)T_b)}\Pi\left(t - \frac{\tau}{2} - (k-1)T_b\right)$$

$$= e^{j2\pi f_k \tau} K_\Pi\left(t - \frac{(2k-1)T_b}{2}, \tau\right).$$

$$\text{(A.9)}$$

From (A.8)-(A.9), it is deduced that the general equation of the autoterms is given as

$$K_{z,\text{auto}}(t, \tau) = \sum_{k=1}^{N} K_{z,k,k}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right), \quad \text{(A.10)}$$

where

$$K_{z,k,k}\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right)$$

$$= P_o \exp(j2\pi f_k \tau) K_\Pi\left(t - \left(\frac{(2k-1)T_b}{2}\right), \tau\right).$$

$$\text{(A.11)}$$

The power of the signal is $P_o = |A_k|^2$ and the rhombic-shaped IAF is centered at $t = (2k-1)T_b$ and $\tau = 0$. From this, it is observed that the autoterms are located along the time axis, centered at lag $\tau = 0$, and have no Doppler frequency, v components. Choosing the lag-window with the size of $-T_b \leq \tau \leq T_b$ will preserve all these terms.

B. CROSS-TERMS

The cross-terms are bilinear products between kth and lth symbols, where $k \neq l$. These are the instantaneous cross-correlation functions (ICFs) of symbols 1, 2, 3, and 4. The

bilinear product between the 1st and the 2nd symbols, where $f_k = f_l = f_1$ are given as

$$K_{z,1,2}(t - T_b, \tau + T_b)$$

$$= e^{j2\pi f_1(t+\tau/2)}\Pi\left(t + \frac{\tau}{2}\right)$$

$$\times e^{-j2\pi f_1(t-\tau/2-T_b)}\Pi\left(t - \frac{\tau}{2} - T_b\right)$$

$$= e^{j2\pi f_1(\tau-T_b)}K_{\Pi}(t - T_b, \tau + T_b),$$

$$K_{z,2,1}(t - T_b, \tau - T_b)$$

$$= e^{j2\pi f_1(t+\tau/2-T_b)}\Pi\left(t + \frac{\tau}{2} - T_b\right)$$

$$\times e^{-j2\pi f_1(t-\tau/2)}\Pi\left(t - \frac{\tau}{2}\right)$$

$$= e^{j2\pi f_1(\tau-T_b)}K_{\Pi}(t - T_b, \tau - T_b).$$

(B.1)

For an arbitrary kth and lth symbols, where $f_k = f_l$, the bilinear product is given as

$$K_{z,k,l}\left(t - \frac{(k+l-1)T_b}{2}, \tau\right)$$

$$= e^{j2\pi f_k(t+\tau/2-(k-1)T_b)}\Pi\left(t + \frac{\tau}{2} - (k-1)T_b\right)$$

$$\times e^{-j2\pi f_k(t-\tau/2-(l-1)T_b)}\Pi\left(t - \frac{\tau}{2} - (l-1)T_b\right)$$

(B.2)

$$= e^{j2\pi f_k \tau}e^{j2\pi((l-k)f_k)T_b}$$

$$\times K_{\Pi}\left(t - \frac{(k+l-1)T_b}{2}, \tau - (k-l)T_b\right).$$

It is observed that these cross-terms have delayed lag components and are located at $\tau > 0$. These terms do not have time component.

On the other hand, the bilinear products between kth and lth symbols, where $f_k \neq f_l$ have both time and lag components. The bilinear products are given as

$$K_{z,1,3}\left(t - \frac{3T_b}{2}, \tau - 2T_b\right)$$

$$= e^{j2\pi f_1(t+\tau/2)}\Pi\left(t + \frac{\tau}{2}\right)$$

$$\times e^{-j2\pi f_0(t-\tau/2-2T_b)}\Pi\left(t - \frac{\tau}{2} - 2T_b\right)$$

$$= e^{j2\pi(f_1-f_0)t}e^{j2\pi(f_1+f_0)\tau/2}e^{j2\pi f_0(2T_b)}$$

$$\times K_{\Pi}\left(t - \frac{3T_b}{2}, \tau - 2T_b\right),$$

$$K_{z,2,3}(t - T_b, \tau + T_b)$$

$$= e^{j2\pi f_1(t+\tau/2-T_b)}\Pi\left(t + \frac{\tau}{2} - T_b\right)$$

$$\times e^{-j2\pi f_0(t-\tau/2-2T_b)}\Pi\left(t - \frac{\tau}{2} - 2T_b\right)$$

$$= e^{j2\pi(f_1-f_0)t}e^{j2\pi(f_1+f_0)\tau/2}e^{j2\pi(2f_0-f_1)T_b}$$

$$\times K_{\Pi}(t - 2T_b, \tau + T_b).$$

(B.3)

It is observed that these cross-terms have both time and lag components and are located away from the time axis. For any kth and lth symbol, where $f_k \neq f_l$, the ICF is

$$K_{z,k,l}\left(t - \frac{(k+l-1)T_b}{2}, \tau\right)$$

$$= e^{j2\pi f_k(t+\tau/2-(k-1)T_b)}\Pi\left(t + \frac{\tau}{2} - (k-1)T_b\right)$$

$$\times e^{-j2\pi f_l(t-\tau/2-(l-1)T_b)}\Pi\left(t - \frac{\tau}{2} - (l-1)T_b\right)$$

(B.4)

$$= e^{j2\pi(f_k-f_l)t}e^{j2\pi(f_1+f_0)\tau/2}e^{j2\pi((l-1)f_l-(k-1)f_k)T_b}$$

$$\times K_{\Pi}\left(t - \frac{(k+l-1)T_b}{2}, \tau - (k-l)T_b\right).$$

From (B.1)–(B.4), the general equation of the cross-terms can be deduced as

$$K_{z,\text{cross}}(t, \tau) = \sum_{\substack{k=0, \\ k \neq l}}^{N-1}\sum_{l=0}^{N-1}K_{z,k,l}\left(t - \frac{(k+l-1)}{2}T_b, \tau - (k-l)T_b\right),$$

(B.5)

where

$$K_{z,k,l}\left(t - \left(\frac{(k+l-1)T_b}{2}\right), \tau - (k-l)T_b\right)$$

$$= P_o \exp\left(j2\pi((k-1)f_k - (l-1)f_l)T_b\right)$$

$$\times \exp\left(j2\pi\left(\frac{(f_k+f_l)}{2}\right)\tau\right)$$

(B.6)

$$\times \exp\left(j2\pi(f_l-f_k)t\right)$$

$$\times K_{\Pi}\left(t - \left(\frac{(k+l-1)T_b}{2}\right), \tau - (k-l)T_b\right).$$

This rhombic-shaped ICF which has the power of $P_o = A_k A_l^*$ is centered at $t = (k+l-1)T_b$ and $\tau = (k-l)T_b$. This shows that the cross-terms are located away from the time axis and the Doppler-frequency axis. In general, the cross-terms have both time and lag components. However, the cross-terms between symbols with the same subcarrier frequency do not have any time components. To remove these cross-terms, low-pass filters with suitable cutoff in lag domain and Doppler-frequency domain must be used.

REFERENCES

[1] M. J. Bastiaans, T. Alieva, and L. Stankovic, "On rotated time-frequency kernels," *IEEE Signal Processing Letters*, vol. 9, no. 11, pp. 378–381, 2002.

[2] D. L. Jones and T. W. Parks, "A high resolution data-adaptive time-frequency representation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 12, pp. 2127–2135, 1990.

[3] D. L. Jones and R. G. Baraniuk, "Simple scheme for adapting time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3530–3535, 1994.

[4] G. R. Arce and S. R. Hasan, "Elimination of interference terms of the discrete Wigner distribution using nonlinear filtering," *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2321–2331, 2000.

[5] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 41, no. 4, pp. 1589–1602, 1993.

[6] R. G. Baraniuk and D. L. Jones, "Signal-dependent time-frequency analysis using a radially Gaussian kernel," *Signal Processing*, vol. 32, no. 3, pp. 263–284, 1993.

[7] S. Krishnamachari and W. J. Williams, "Adaptive kernel design in the generalized marginals domain for time-frequency analysis," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '94)*, vol. 3, pp. 341–344, Adelaide, Australia, April 1994.

[8] D. Boutana, B. Barkat, and F. Marir, "A proposed high-resolution time-frequency distribution for the analysis of multicomponent and speech signals," *International Journal of Signal Processing*, vol. 1, no. 3, pp. 164–167, 2005.

[9] B. Boashash, "Heuristic formulation of time-frequency distributions," in *Time Frequency Signal Analysis and Processing: A Comprehensive Reference*, B. Boashash, Ed., pp. 29–57, Elsevier, Amsterdam, The Netherlands, 2003.

[10] L. Stankovic, "Autoterm representation by the reduced interference distributions: a procedure for kernel design," *IEEE Transactions on Signal Processing*, vol. 44, no. 6, pp. 1557–1563, 1996.

[11] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Transactions on Signal Processing*, vol. 40, no. 2, pp. 402–412, 1992.

[12] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1253–1263, 2003.

[13] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—I: fundamentals," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520–538, 1992.

[14] L. Stankovic, M. Dakovic, J. Jiang, and E. Sejdic, "Instantaneous frequency estimation using the *S*-transform," *IEEE Signal Processing Letters*, vol. 15, pp. 309–312, 2008.

[15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, 2nd edition, 1992.

[16] B. Sklar, *Digital Communications: Fundamentals and Applications*, Prentice-Hall, Upper Saddle River, NJ, USA, 2nd edition, 2001.

[17] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.

[18] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, Upper Saddle River, NJ, USA, 2000.

[19] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Transactions on Communications*, vol. 46, no. 4, pp. 431–436, 1998.

[20] M. Paprzycki and P. Stpiczyński, "A brief introduction to parallel computing," in *Handbook of Parallel Computing and Statistics*, E. J. Kontoghiorghes, Ed., pp. 3–41, Chapman & Hall, Boca Raton, Fla, USA, 2006.

[21] J. Fan and J. S. Marron, "Fast implementations of nonparametric curve estimators," *Journal of Computational and Graphical Statistics*, vol. 3, no. 1, pp. 35–56, 1994.

[22] W. S. Cleveland and E. Grosse, "Computational methods for local regression," *Statistics and Computing*, vol. 1, no. 1, pp. 47–62, 1991.