ABSTRACT
This paper addresses abrupt change detection in multiplicative noise using the continuous wavelet transform. An optimal wavelet, maximizing a well-chosen time-scale contrast criterion is derived. The analytical optimization gives the optimal wavelet closed expression. The influence of the mother wavelet on signature-based detector performance is then demonstrated. Detection performance is characterized using Receiver Operating Characteristic curves computed from Monte-Carlo simulations. The optimal wavelet obviously improves performance with respect to other wavelets classically used for singularity detection.

1. INTRODUCTION AND PROBLEM STATEMENT
Object contour extraction or signal segmentation require the detection of Abrupt Changes (AC) in the parameters of the observed signals or images [1]. Additive noise models have received considerable attention for this segmentation problem. This paper addresses the problem of detecting AC in the parameters of signals/images corrupted by multiplicative noise. Multiplicative noise models have been used successfully in many applications. These applications include image segmentation in systems using coherent radiation (radar, laser) [2]. The Continuous Wavelet Transform (CWT) has shown interesting properties for AC detection in additive and/or multiplicative noise models [3]. Indeed, the CWT provides an AC signature emerging from the transformed noise as scale increases. This paper derives an optimal wavelet for AC detection in multiplicative noise models. This optimal wavelet maximizes a performance criterion expressed as a time-scale contrast. It is interesting to note that the problem of wavelet optimization has already been considered in [7] for classification purposes. The criterion to be optimized was expressed as a function of the classification error for this application. However, the optimization was based on a particular parametrization of the wavelet-associated scaling filter. This paper proposes an analytical optimization yielding a closed form expression of the optimal wavelet. This expression does not require a priori parametric model for the wavelet as in [7].

The optimal wavelet is then studied for AC detection in multiplicative noise models. Sum across scales of the CWT has shown interesting properties for AC detection in [3]. This paper shows that the performance of the AC detector is strongly related to the choice of the mother wavelet. The optimal wavelet yields the best performance compared to other wavelets classically used for singularity detection.

Denote as \( x(t) = m_x + b(t) \) the multiplicative noise, where \( b(t) \) is a stationary white noise with power spectral density \( N_x \) and \( m_x = E[x(t)] \neq 0 \) (the multiplicative noise is usually modeled as a non-zero mean process). The AC is modeled as follows:

\[
s(t) = 1 + AU(t - t_0) \quad t \in \Omega, \ t_0 \in \Omega, A \geq 0,
\]

where \( U(t) \) denotes an ideal step, \( \Omega \) is the observation interval, \( t_0 \) is the AC location and \( s(t) \) models an ideal AC whose amplitude varies from 1 to \( 1 + A \). The AC detection for a multiplicative noise model (studied in [2] for edge detection in synthetic aperture radar images) expresses as the following simple binary hypothesis test [10]:

\[
H_0 : y(t) = x(t),
H_1 : y(t) = x(t)s(t).
\]

The AC detection problem can be expressed in the time-scale domain using the CWT. The CWT of \( y(t) \) is defined by:

\[
C_y(a, \tau) = \int_{-\infty}^{+\infty} y(t) \psi^*_a,\tau(t) dt, \quad a \in \mathbb{R}^+, \ \tau \in \mathbb{R}
\]

with \( \psi_{a,\tau} = \frac{1}{\sqrt{a}} \psi \left( \frac{t - \tau}{a} \right) \).

The analyzing function family \( \{ \psi_{a,\tau} \}_{a \in \mathbb{R}^+, \ \tau \in \mathbb{R}} \) is obtained by dilation and translation of a function \( \psi \) called the mother wavelet (\( a \) is the scale and \( \tau \) is the translation parameter). If \( \psi \) satisfies the admissibility condition (which is \( \int_{-\infty}^{+\infty} \psi(t) dt = 0 \) when the Fourier transform of \( \psi \) is continuous), the transform has a reconstruction formula [4]. This study is restricted to real normalized wavelets with bounded support \([\Delta_1, \Delta_2]\). Time-scale detection considers the following simple binary hypothesis test:

\[
H_0 : C_y(a, \tau) = \int_{-\infty}^{+\infty} x(t) \psi^*_{a,\tau}(t) dt,
H_1 : C_y(a, \tau) = \int_{-\infty}^{+\infty} x(t)s(t) \psi^*_{a,\tau}(t) dt,
\]

for \( a \in \mathbb{R}^+, \ \tau \in \mathbb{R} \). A linear time-scale detector based on the sum along scales of the CWT is considered in section 2. This detector does not require a priori knowledge of the noise distribution and is consequently suboptimal. A contrast criterion is a measure of suboptimal detection performance. General contrast definition is presented in section 2. However the choice of an appropriate time-scale contrast is the first critical point for wavelet optimization. CWT moments derivation under both hypothesis in section 2 justifies the choice of the complementary deflection in the multiplicative noise case. Moreover, this contrast increases with scale which demonstrates the interest of working in the time-scale plane. Section 3 derives the optimal wavelet maximizing the contrast criterion under normality and admissibility constraints. ROC curves demonstrate the improvement of signature-based detection performance obtained with this wavelet in section 3.

2. TIME-SCALE DETECTION
2.1. Time-scale detector
The CWT is a correlation between the observed process and a scaled mother wavelet. The two-dimensional thresholding of the
CWT has already shown interesting properties for AC detection [5]. However, the knowledge of the CWT distribution under hypothesis $H_0$ is necessary to adjust the threshold as a function of the probability of false alarm. This distribution may be unknown or difficult to derive when the multiplicative noise $x(t)$ is non-Gaussian. The summation of the CWT over several scale octaves (denoted $\Gamma$) was proposed in [3] to overcome this problem

$$\Gamma(\tau) = \sum_{i=1}^{n} C_y(a_i, \tau) .$$

This summation yields Gaussian test statistics (by central limit theorem) under mild assumptions regarding the multiplicative noise distribution. Moreover, summing over different octaves reduce noise effects since the noise maxima do not propagate from one scale to another [8]. Fig. 1 displays a run of this detector for an observed process of $N = 512$ samples. The AC location and amplitude are respectively $t_0 = 256$ and $A = 0.4$. The multiplicative noise parameters are $\sigma^2_n = 0.2$ and $m_n = 1$. The CWT is approached by a discrete wavelet transform derived for dyadic scales i.e. $a = 1, 2, 4, 8, ..., 512$ and without downsampling in time (i.e. $\tau = 1, 2, 3, ..., 512$ for each scale value) also called dyadic wavelet transform [8]. The detector $\Gamma(\tau)$ clearly shows a maximum when an AC occurs (and two minima due to boundary effects), contrary to the absence of change. The next section defines a time-scale contrast which allows to determine appropriate mother wavelets for AC detection.

### 2.2. Time-scale contrast

#### 2.2.1. General expression

This section defines an appropriate time scale contrast for the detection problem (1). Many contrast criterion have been defined in the literature [6], [9]. However, criteria based on first and second-order moments are often preferred by simplicity. Let $m_i(\alpha, \tau)$ and $\sigma^2_i(\alpha, \tau)$ denote the mean and variance of $C_y(a_i, \tau)$ under hypothesis $H_i, i = 0, 1$. This paper proposes to study the following time-scale contrast:

$$F_\alpha(\alpha, \tau) = \frac{m_1(\alpha, \tau) - m_0(\alpha, \tau)}{\sqrt{\text{var}(C_y(a_i, \tau))}} ,$$

where $\text{var}_n[C_y(a_i, \tau)]$ is the variance corresponding to the mixing distribution $p_n(x) = (1 - \alpha)p_0(x) + \alpha p_1(x), \alpha \in [0, 1]$, $p_0(.)$ and $p_1(.)$ are the distributions of $C_y(a_i, \tau)$ under hypotheses $H_0$ and $H_1$. The criteria obtained for $\alpha = 0$ and $\alpha = 1$ are usually referred to as deflection and complementary deflection respectively [6]. The deflection and complementary deflection are equal when the observed signal $s(t)$ is corrupted by additive noise (variance is the same under both hypotheses). However, this result is no longer valid in the multiplicative noise case. Indeed, the CWT moment derivation performed in the next subsection allows the choice of an appropriate time-scale contrast in the multiplicative noise case.

#### 2.2.2. Complementary deflection

Let $m_i(\alpha, \tau)$ and $\sigma^2_i(\alpha, \tau)$ denote the mean and variance of $C_y(a_i, \tau)$ under hypothesis $H_i, i = 0, 1$. The admissibility condition yields:

$$m_0(a, \tau) = m_n \int_{-\infty}^{+\infty} \psi^2_{\alpha, \tau}(t) dt = 0.$$

Moreover, the CWT is invariant with respect to translation and dilation of the original signal. This property allows to define an AC signature in the time-scale domain defined as the mean of $C_y(a, \tau)$ under hypothesis $H_1$:

$$m_1(a, \tau) = \begin{cases} \frac{m_n \int_{-\infty}^{+\infty} \psi^2_{\alpha, \tau}(t) dt}{m_n \int_{-\infty}^{+\infty} \psi^2_{\alpha, \tau}(t) dt} & \text{for } \frac{m_n}{\var_\alpha} \in [\Delta_1, \Delta_2] \\ 0 & \text{else} \end{cases} .$$

The AC signature is conic and points to the AC instant $t_0$. Moreover, the square modulus of the signature is proportional to the scale on the straight line $D_k$ defined by $\frac{m_n}{\var_\alpha} = \xi$ (with $\xi \in [\Delta_1, \Delta_2]$). Straightforward computations show to determine the variance of $C_y(a, \tau)$ under hypotheses $H_0$ and $H_1$. Under hypothesis $H_0$, the CWT is a centered random field with constant variance:

$$\sigma^2_0(a, \tau) = \sigma^2_\alpha \int_{-\infty}^{+\infty} |\psi_{n, \tau}(t)|^2 dt = \sigma^2_\alpha \int_{-\infty}^{+\infty} |\psi_{n, \tau}(t)|^2 dt = \sigma^2_\alpha .$$

Under hypothesis $H_1$, the AC signature is embedded in a time-scale noise whose variance is defined by:

$$\sigma^2_1(a, \tau) = \sigma^2_\alpha \int_{-\infty}^{+\infty} |\psi_{n, \tau}(t)|^2 dt = \sigma^2_\alpha \int_{-\infty}^{+\infty} |\psi_{n, \tau}(t)|^2 dt = \sigma^2_\alpha .$$

The next subsection derives an expression for the AC detection problem (1). Many contrast criterion have been defined in the literature [6], [9]. However, criteria based on first and second-order moments are often preferred by simplicity. Let $m_i(\alpha, \tau)$ and $\sigma^2_i(\alpha, \tau)$ denote the mean and variance of $C_y(a_i, \tau)$ under hypothesis $H_i, i = 0, 1$. This paper proposes to study the following time-scale contrast:

$$F_\alpha(\alpha, \tau) = \frac{m_1(\alpha, \tau) - m_0(\alpha, \tau)}{\sqrt{\text{var}(C_y(a_i, \tau))}} ,$$

for $\xi \in [\Delta_1, \Delta_2]$. This contrast is proportional to the scale on each line $D_k$ of the time-scale domain (defined by $\xi = \text{constant}$). This highlights the interest of working in the time-scale domain. Denote $F(\psi, \xi)$ as the time-scale contrast on the straight line $D_k$ (i.e. $F(\psi, \xi) = F_1(a, t_0 - a\xi)$. The next subsection derives an optimal wavelet $\psi$ maximizing $F(\psi, \xi)$ under appropriate normalization and admissibility conditions.

### 3. Optimal Wavelet

The contrast $F(\psi, \xi)$ is maximized on the set of normalized and admissible wavelets with bounded support $[\Delta_1, \Delta_2]$. Normalization and admissibility conditions expressed as:

$$\int_{\Delta_1}^{\Delta_2} |\psi(t)|^2 dt = 1.$$
A wavelet $\psi(t)$ is said optimal if it maximizes the complementary deflection under constraints (6) and (7). The contrast criterion evaluated on the straight line $D^2$ is proportional to the scale $a$. Its maximization with respect to the mother wavelet $\psi$ for a fixed value of $\xi$ is equivalent to the maximization of:

\[ F(\psi, \xi) = \frac{|P_{\psi, \xi}|^2}{1 + \nu N_{\psi, \xi}^2} + \frac{N_{\psi, \xi}^2 (\Delta_2 - \xi)}{P_{\psi, \xi}} \int_{\Delta_1}^{\Delta_2} \psi(t) dt = 0. \quad (7) \]

The normalization constraint (6) clearly yields $0 \leq N_{\psi, \xi}^2 \leq 1$. The Cauchy-Schwarz inequality applied to $|P_{\psi, \xi}|^2$ can be written as

\[ |P_{\psi, \xi}|^2 \leq N_{\psi, \xi}^2 (\Delta_2 - \xi). \quad (8) \]

Constraint (7) implies that

\[ \int_{\Delta_1}^{\Delta_2} \psi(t) dt = -P_{\psi, \xi}. \]

The Cauchy-Schwarz inequality in $L^2 ([\Delta_1, \xi])$ yields:

\[ |P_{\psi, \xi}|^2 \leq (\xi - \Delta_1) \int_{\Delta_1}^{\Delta_2} |\psi(t)|^2 dt = (\xi - \Delta_1) (1 - N_{\psi, \xi}^2). \quad (9) \]

Fig. 2 displays the authorized variation domain of $([P_{\psi, \xi}|^2, N_{\psi, \xi}^2]$ imposed by (8) and (9). This variation domain is a triangular region delimited by the two following straight lines:

\[ |P_{\psi, \xi}|^2 = N_{\psi, \xi}^2 (\Delta_2 - \xi) \quad (10) \]

\[ |P_{\psi, \xi}|^2 = (\xi - \Delta_1) (1 - N_{\psi, \xi}^2) \quad (11) \]

These two straight line intersect at the point defined by:

\[ N_{\psi, \xi}^2 = \frac{(\xi - \Delta_1)}{(\Delta_2 - \xi)}. \]

Maximizing $F(\psi, \xi)$ for a given $\xi$ in $[\Delta_1, \Delta_2]$ and a given $N_{\psi, \xi}^2$ in $[0, 1]$ is equivalent to maximize $|P_{\psi, \xi}|^2$.

Let $|P_{\psi, \xi}|^2$ denote the value of $|P_{\psi, \xi}|^2$ leading to a maximum contrast for given $\xi$ and $N_{\psi, \xi}^2$. $|P_{\psi, \xi}|^2$ is obtained when Cauchy-Schwarz inequalities given by (8) and (9) become equalities i.e. for $([P_{\psi, \xi}|^2, N_{\psi, \xi}^2]$ on the boundary segments defined by:

\[ \left| P_{\psi, \xi} \right|^2 = N_{\psi, \xi}^2 (\Delta_2 - \xi), \quad N_{\psi, \xi}^2 \in [0, N_{\psi, \xi}^2], \quad N_{\psi, \xi}^2 \in [N_{\psi, \xi}^2, 1]. \quad (12) \]

Denote $\psi|_{[\Delta_1, \xi]}$ (respectively $\psi|_{[\xi, \Delta_2]}$) as the restriction of $\psi$ to the interval $[\Delta_1, \xi]$ (respectively $[\xi, \Delta_2]$). The Cauchy-Schwarz inequality (8) (respectively (9)) becomes an equality if $\psi|_{[\xi, \Delta_2]}$ (respectively $\psi|_{[\Delta_1, \xi]}$) is colinear to the constant function $1_{[\xi, \Delta_2]}$ (respectively $1_{[\Delta_1, \xi]}$). Consequently, the optimal wavelet maximizing $F(\psi, \xi)$ (for given $\xi \in [\Delta_1, \Delta_2]$ and a given $N_{\psi, \xi}^2 \in [0, 1]$) is defined by:

\[ \psi = C_1 1_{[\Delta_1, \xi]} + C_2 1_{[\xi, \Delta_2]}. \]

The constants $C_1$ and $C_2$ can be derived from (12) as functions of $N_{\psi, \xi}^2$ and $\xi$:

\[ |C_2|^2 = \frac{N_{\psi, \xi}^2}{\Delta_2 - \xi} \quad \text{and} \quad C_1 = \frac{(\Delta_2 - \xi)(\Delta_1 - \xi)}{(\Delta_2 - \xi)(\Delta_1 - \xi)} \]

Consequently, the contrast expresses as:

\[ F(\psi, \xi) = \frac{N_{\psi, \xi}^2 (\Delta_2 - \xi)}{1 + \nu N_{\psi, \xi}^2} \quad \text{for} \quad N_{\psi, \xi}^2 \in [0, N_{\psi, \xi}^2]. \]

\[ F(\psi, \xi) = \frac{(\xi - \Delta_1)(1 - N_{\psi, \xi}^2)}{1 + \nu N_{\psi, \xi}^2} \quad \text{for} \quad N_{\psi, \xi}^2 \in [N_{\psi, \xi}^2, 1]. \]

Straightforward computations show that $F(\psi, \xi)$ is an increasing (respectively decreasing) function of $N_{\psi, \xi}^2$ over $[0, N_{\psi, \xi}^2]$ (respectively $[N_{\psi, \xi}^2, 1]$). The maximum contrast value is then obtained for $([P_{\psi, \xi}|^2, N_{\psi, \xi}^2]$ at the intersection of the two segments i.e. for $N_{\psi, \xi}^2 = N_{\psi, \xi}^2$. The maximum contrast value for a given $\xi$ in $[\Delta_1, \Delta_2]$ is given by:

\[ F(\xi) = \frac{(\Delta_2 - \xi)(\xi - \Delta_1)}{\Delta_2 - \Delta_1 + \nu(\xi - \Delta_1)} \]

Note that the wavelet maximizing the contrast $F(\psi, \xi)$ has been determined for a fixed support $[\Delta_1, \Delta_2]$ and for each value of $\xi \in [\Delta_1, \Delta_2]$. The value of $\xi \in [\Delta_1, \Delta_2]$, denoted $\xi_{\text{opt}}$, for which the contrast is maximal with respect to $\xi$ can be determined by derivation of $F(\xi)$ with respect to $\xi$:

\[ \xi_{\text{opt}} = \Delta_1 + \frac{\Delta_2 - \Delta_1}{A + 2} \quad (13) \]

Finally, for a given AC amplitude, the optimal wavelet $\psi_{\text{opt}}$, is defined by:

\[ \left\{ \begin{array}{l}
\psi_{\text{opt}}(t) = C_1 = -\xi \sqrt{\frac{(A + 1)}{(A + 2)(\Delta_2 - \Delta_1)}} \quad t \in [\Delta_1, \xi_{\text{opt}}] \\
\psi_{\text{opt}}(t) = C_2 = \frac{1}{\sqrt{(A + 1)(\Delta_2 - \Delta_1)}} \quad t \in [\xi_{\text{opt}}, \Delta_2]
\end{array} \right. \]

with $\xi_{\text{opt}}$ defined by (13). The contrast is maximum on the line $D_{\text{opt}}$ of the time-scale plane. The optimal wavelet is a distorted version of Haar wavelet and is function of the AC amplitude. When no a priori knowledge of the AC amplitude is available, the optimal wavelet can be approached by the Haar wavelet for small AC amplitude in the multiplicative noise case. Indeed, for $A \ll 1$:

\[ \xi_{\text{opt}} \cong \frac{\Delta_1 + \Delta_2}{2} \]

\[ C_1 \cong -\frac{1}{\sqrt{(A + 1)(\Delta_2 - \Delta_1)}} \quad \text{and} \quad C_2 \cong \frac{1}{\sqrt{(A + 1)(\Delta_2 - \Delta_1)}} \]

Moreover, the ratio between maximal contrasts obtained with the Haar wavelet and the optimal wavelet is such that:

\[ \frac{F(\psi_{\text{Haar}, \xi_{\text{opt}}})}{F(\psi_{\text{opt}}, \xi_{\text{opt}})} = \frac{(A + 2)^2}{2(1 + (1 + A)^2)} \cong 1 \quad \text{for} \quad A \ll 1 \]

4. SIMULATION RESULTS

Many simulations have been conducted to determine how detector performance depends on the mother wavelet. In this paper, the
parameters are $N = 512$ (number of samples), $t_0 = 256$ (AC location), $A = 0.4$ (AC amplitude), $\sigma^2 = 1$ (multiplicative noise variance) and $m_0 = 1$ (multiplicative noise mean value). The performance of the detector 2 is compared for different mother wavelets: the optimal wavelet, the Haar wavelet and the first and second derivatives of a Gaussian wavelets. The CWT is derived for the following scales $a = 1, 2, 4, 8, ..., 256, 512$. First and second derivative of a Gaussian wavelets (the second is also referred to as Mexican Hat) are frequently used for singularity characterization or detection. Receiver Operating Characteristics (ROCs) have been estimated from 10000 Monte-Carlo runs by comparing the test statistic to 500 different threshold values. As shown in Fig. 3, the detection performance is better with the optimal wavelet than with the first and second derivative of a Gaussian wavelets. However, for small AC amplitudes, the optimal wavelet can be approximated by the Haar wavelet. Indeed, the probability of detection (PD) for the optimal and the Haar wavelet for a fixed probability of false alarm (PFA).

5. CONCLUSION

This paper analytically derived an optimal wavelet for AC detection in multiplicative noise. This optimal wavelet was obtained by maximizing an appropriate time-scale contrast under normality and admissibility constraints. Simulation results show a significant decrease in AC detection performance when the optimal wavelet is not used.

Other contrasts based on pre- and/or post-processings of the CWT are currently under consideration. The possible processings include the weighted sum of the scalogram or time-scale correlation with an a priori known signature [3].

6. REFERENCES