Adaptive wavelet transform with application in signal denoising

Tomić, Mladen

UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Mladen Tomić

ADAPTIVE WAVELET TRANSFORM WITH APPLICATION IN SIGNAL DENOISING

ADAPTIVNA VALIĆNA TRANSFORMACIJA S PRIMJENOM NA UKLANJANJE ŠUMA IZ SIGNALA

DOCTORAL THESIS
DOKTORSKA DISERTACIJA

Zagreb, 2010

The doctoral thesis has been realized at the Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia, and at the Faculty of Engineering, University of Rijeka, Croatia.

Supervisor: Dr.sc. Damir Seršić, associate professor

The doctoral thesis consists of 155 pages

Doctoral thesis number: _____

The dissertation evaluation committee:

1. Professor Sven Lončarić, Ph.D.,
   Faculty of Electrical Engineering and Computing, University of Zagreb
2. Professor Damir Seršić, Ph.D.,
   Faculty of Electrical Engineering and Computing, University of Zagreb
3. Assistant Professor Viktor Sučić, Ph.D.,
   Faculty of Engineering, University of Rijeka

The dissertation defense committee:

1. Professor Sven Lončarić, Ph.D.,
   Faculty of Electrical Engineering and Computing, University of Zagreb
2. Professor Damir Seršić, Ph.D.,
   Faculty of Electrical Engineering and Computing, University of Zagreb
3. Assistant Professor Viktor Sučić, Ph.D.,
   Faculty of Engineering, University of Rijeka
4. Assistant Professor Miroslav Vrankić, Ph.D.,
   Faculty of Engineering, University of Rijeka
5. Professor Sonja Grgić, Ph.D.,
   Faculty of Electrical Engineering and Computing, University of Zagreb

Date of dissertation defense: 26^{th} February 2010

Acknowledgments

It is a pleasure to express my appreciation to my supervisor, Professor Damir Seršić, who always expressed a spirit of quest in regard to research. His engagement that included a supportive and friendly approach made the thesis research a nice experience.

I would like to thank Professor Miroslav Vrankić who introduced me to the field of signal denoising and whose endless assistance in my initial research had lasting effect.

I am indebted to thank my colleagues from the Faculty of engineering who were giving me support and motivation to expedite and finish the research.

I would not be who I am without my family. A special thank you goes to my parents who have throughout my life stood by me and whose continued support has lasted to date.

I lack words to describe joy and happiness which my wife and kids gave me every day. Their immense love and faith continuously motivated me throughout this long and bumpy journey.

Most importantly, I am grateful for having God's love in my life, as He has always been good to me, guiding me through life; He offered a helping hand when needed and carried me and my family during the most challenging of times.

Contents

Chapter 1

Introduction

1.1 Motivation

Most of the real-world signals are corrupted by noise, but noise level is sometimes too low to cause concern for a given application. When it is not the case, it is necessary to employ efficient signal denoising algorithms to improve signal quality. Wavelet transforms were proved as a valuable tool in many digital signal processing applications, including signal denoising.

Although wavelet theory has been known for a long time, their application in signal processing became widespread only after Mallat successfully linked the theory to the filter bank theory (Mallat, 1988), showing that the wavelet transform can be easily calculated by simple signal filtering. First, a low-pass and a high-pass filters are used on original signal, and then the filtering step is iterated over the low-pass branch. The final results are a low-pass subband containing only a coarse signal information, and high-pass subbands, containing successively finer and finer signal details.

Compared to other well known transforms, such as the Fourier transform, compactly supported wavelets have advantage that they achieve good localization of both the time and the frequency. Because of the compact support, wavelets can much better represent non-periodic and non-stationary signals, especially the ones containing sharp changes in local properties, such as discontinuities. Another important property of wavelets is sparseness of signal representation in the transform domain. It means that most of the signal information is contained in a small number of larger magnitude wavelet coefficients. Furthermore, white noise in a signal is, in the transform domain, represented by a large number of smaller magnitude wavelet coefficients, concentrated about zero. The two properties lead to a conclusion that signal denoising can be efficiently accomplished by simple wavelet coefficients thresholding. A brief summary of wavelet transforms in signal denoising is given in (Tomic, 2008a,b).

Popularity of wavelet transforms was further increased with the work of Sweldens, who proposed ladder structure for the transform calculation, called the lifting scheme (Sweldens, 1996). The benefits of the new approach are that it allows for more efficient

transform calculation, makes the perfect reconstruction property guaranteed, performs all calculations in the spatial domain and makes it easy to introduce adaptivity to a transform. Wavelet construction is now also being performed in the spatial domain, instead of the frequency domain, as in the classical approach. It became possible to construct wavelets which are not necessarily dilates and translates of a single mother wavelet function. They allow for efficient handling of signal boundaries, irregularly sampled signals or certain signal features. Since they inherit all the good properties of classical wavelets (first generation wavelets), but also introduce new features and advantages, they are called the second generation wavelets.

In a separate field, Katkovnik proposed successful denoising methods built upon local polynomial approximation (LPA) and intersection of confidence intervals (ICI) rule. The ICI rule was used as a tool for adaptive support selection. First the 1-D case was examined (Katkovnik, 1999) and afterwards the ICI was used for very efficient image denoising algorithms (Foi, Katkovnik, and Egiazarian, 2007; Katkovnik, Egiazarian, and Astola, 2003). The ICI rule is also employed as an adaptation method in the wavelet based denoising algorithm proposed in this thesis.

1.2 Improving Denoising Performance of Wavelet Transforms

Most denoising algorithms, including wavelet transforms, generally perform very well for smooth signals or smooth signal regions. The greatest difference in the algorithm efficiency can be noted for signals containing higher frequency features, such as discontinuities, spikes, bumps, or any other sudden change in local signal properties. Although compactly supported wavelets stand out as an excellent performer for such signal classes, there is still much room for improvements.

A common approach to wavelet transforms adaptation is introduction of algorithms for basis functions selection. Based on the adaptivity criteria, one of wavelets from a predefined set of wavelet functions is chosen. The selection is carried out in order to find the optimal wavelet for a whole signal, for each decomposition scale or for each individual point. Longer and smoother wavelets can efficiently represent smooth signals or signal regions, while shorter or irregular wavelets can efficiently describe higher frequency features in a given signal. Wavelet selection guidelines are, thus, straightforward. Basically, a wavelet which shape best fits signal features should result in the highest quality denoised signal.

We propose an adaptive algorithm which selects wavelet basis function on a point-by-point basis. The ICI rule is used as the core of the adaptive algorithm. Should the adaptive algorithm succeed, for each point, a wavelet shall be chosen which best matches local signal properties in the points neighborhood.

1.3 Related Work

Many adaptive wavelet transforms were proposed which try to improve performance in higher frequency signal regions. They are mostly based on the lifting concept and try to follow the basic guideline for wavelet basis selection.

Claypoole et al. (Claypoole, Baraniuk, and Nowak, 1998; Claypoole, Davis, Sweldens, and Baraniuk, 1997, 2003) proposed two adaptive transforms - the scale-adaptive transform (ScAT) and the space-adaptive transform (SpAT). In the ScAT, an adaptation is performed on a scale-by-scale basis. From a set of predictors they choose the one which better fits the prevailing signal properties. The objective measure of the quality of fit is a simple sum of squared prediction errors. The predictor which yields the smallest sum of squared errors is chosen as the predictor for the entire signal at a given scale.

The approach in the SpAT transform is to perform adaptation on a point-by-point basis. The predictor is chosen for each data point as the one which gives the smallest prediction error. It is to be expected that such predictor will often fit well to the local signal properties and make it possible for the wavelets to contract in the neighborhood of signal discontinuities. The shortcoming of SpAT is that it makes the predictor space-varying. Since the update step depends on the predictor it may complicate its design or introduce the necessity of bookkeeping of the predictor selection. To resolve this, Claypoole proposed the update first framework where the update step is performed before the predict step.

Unlike Claypoole et al. who adapt the predictor, Piella et al. in (Piella and Heijmans, 2002) proposed an adaptive update transform where only the update filter is adapted, while the predictor is fixed. In the general form presented, choice of the update filter depends on the decision maps which can be based on any meaningful criteria. In the original work, an example of a gradient based decision maps is suggested, while in (Piella, Pesquet-Popescu, and Heijmans, 2002) derivative based decision maps are proposed. However, the method is only briefly investigated and no comparison to other methods are provided.

Wu et al. in (Wu, Pan, Zhang, and Zhang, 2004) proposed the algorithm called the switch thresholding. It is a very simple adaptation method in which two conventional wavelets are used - Haar and CDF(2,2). The adaptation is performed on a point-by-point basis by choosing the appropriate wavelet for the current point. For the step edges, Haar transform is to be used, while for the other parts of the signal CDF(2,2) is chosen. Whether the current point is part of a step edge is determined by the difference to its neighboring point. If the difference is larger than the threshold it is considered to be a step edge. Although some coarse limits for the threshold are suggested to follow from the statistics theory, the actual threshold value is determined empirically.

Instead of modifying the transform Chan and Zhou (Chan and Zhou, 1999) took the opposite approach. They showed that the signal itself can be modified in the place of discontinuities. After the discontinuity is detected, signal values from one side of

the discontinuity is used to extrapolate its values to the other side of the discontinuity. Traditional wavelets may now be applied to such a signal region because the discontinuity is eliminated and the whole region is smooth. If we record how the changes are made to the signal, it should be possible to recover the discontinuities at synthesis, using the inverse filters.

Finally, similar approach to the one taken in this thesis was explored by Sersic in (Sersic, 2000a,b,c). He proposed a two-channel wavelet filter bank with adaptive number of zero moments. His basic idea was to use filter banks with more zero moments for smoother parts of a signal, while for the transients and singularities, filter banks with less zero moments were to be used. In (Sersic and Vrankic, 2002), the idea was also investigated for the 2-D wavelet filter banks.

Chapter 2

Wavelet Transform and Wavelet Filter Banks

A brief wavelet theory overview is given in this Chapter. We will introduce multiresolution analysis, scaling and wavelet functions, and present the lifting scheme concept, which is used in the thesis as a wavelet transform implementation tool.

2.1 Multiresolution Analysis

Theory of multiresolution analysis is central to a wavelet transform understanding. It allows for signals to be decomposed into finer and finer details, allowing its analysis and processing at several different resolutions (Burrus, Gopinath, and Guo, 1998). To present the concept, let us first define a set of scaling functions, each being a translate of the basic scaling function, given by:

$$\varphi_k(t) = \varphi(t-k) \qquad k \in \mathbb{Z} \qquad \varphi \in L^2. \tag{2.1}$$

The functions span a subspace of $L^2(\mathbb{R})$, given by:

$$\nu_0 = \overline{\underset{k}{Span}\{\varphi_k(t)\}}, \tag{2.2}$$

Functions $f(t)$ belonging to the space ν_0, can now be analyzed using linear decomposition:

$$f(t) = \sum_k a_k \varphi_k(t), \qquad f(t) \in \nu_0 \tag{2.3}$$

Instead of only translating the basic scaling function (2.1), and staying confined to ν_0, it is also possible to increase the size of the subspace spanned, by changing the scaling functions time scale:

$$\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k), \tag{2.4}$$

Subspace spanned by the functions (2.3) is:

$$\nu_j = \overline{\underset{k}{Span}\{\varphi_k(2^j t)\}} = \overline{\underset{k}{Span}\{\varphi_{j,k}(t)\}}. \tag{2.5}$$

As a consequence, any $f(t) \in \nu_j$, can be expressed using the following decomposition:

$$f(t) = \sum_k a_k \varphi(2^j t + k). \tag{2.6}$$

Scale parameter j defines the resolution at which a signal is analyzed. Spanned subspace grows for $j > 0$, since $\varphi_{j,k}(t)$ is narrower and can represent finer details. For $j < 0$, the span is becoming smaller. Because $\varphi_{j,k}(t)$ is becoming wider and translated in larger steps, it can only represent coarse signal information. The two extreme cases for j are:

$$\begin{aligned} \nu_{-\infty} &= \{0\}, \\ \nu_{\infty} &= L^2. \end{aligned} \tag{2.7}$$

Generally,

$$\begin{aligned} \nu_j \subset \nu_{j+1} \quad \forall j \in \mathbb{Z}, \\ \{0\} \cdots \subset \nu_{-2} \subset \nu_{-1} \subset \nu_0 \subset \nu_1 \subset \nu_2 \subset \cdots L^2. \end{aligned} \tag{2.8}$$

Since ν_0 is a subspace of ν_1, function $\varphi(t)$, belonging to ν_0, also belongs to ν_1. As such, $\varphi(t)$ can be written as a linear combination of basis functions from ν_1. It leads to a dilation (or refinement) equation, which characterizes a multiresolution analysis:

$$\varphi(t) = \sqrt{2} \sum_n h(n) \varphi(2t - n), \qquad n \in \mathbb{Z}, \tag{2.9}$$

where weights $h(n)$ are real or complex numbers, called scaling coefficients, while $\sqrt{2}$ maintains the norm of the scaling function.

2.2 Wavelet Functions

To better describe important features of the signal, we introduce a set of functions $\psi_{j,k}(t)$, called wavelet functions. Unlike the scaling functions $\varphi_{j,k}(t)$, which span spaces ν_j, wavelet functions span differences between the spaces. Let $\mathcal{W}_j$ be the orthogonal complement of ν_j in ν_{j+1}. It means that any subspace ν_{j+1} can be written as a direct sum:

$$\nu_{j+1} = \nu_j \oplus \mathcal{W}_0. \tag{2.10}$$

In general, we can define the whole L^2 by starting from an initial arbitrarily chosen

subspace ν_{j_0} and successively adding the differences spanned by the wavelet functions:

$$L^2 = \nu_{j_0} \oplus \mathcal{W}_{j_0} \oplus \mathcal{W}_{j_0+1} \oplus \mathcal{W}_{j_0+2} \oplus \cdots . \tag{2.11}$$

From the signal processing point of view, it is advantageous to choose the initial scale j_0 such that it represents coarse signal information. The coarse information is essential for signal reconstruction and it is most often fully preserved. Subspaces $\mathcal{W}_j$ provide finer and finer details which can be analyzed or processed to achieve the desired goal.

Since the wavelet functions $\psi_{j,k}(t)$ reside in the subspace ν_{j+1}, spanned by the next narrower scaling functions $\varphi_{j+1,k}(t)$, they can also be expressed as their linear combination:

$$\psi(t) = \sqrt{2} \sum_n h_1(n) \varphi(2t - n), \qquad n \in \mathbb{Z}, \tag{2.12}$$

where $h_1(n)$ are wavelet coefficients. The function $\psi(t)$ obtained by the above expression is called the mother wavelet. All the other wavelet functions that form the expansion basis, are translated and scaled versions of the mother wavelet:

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k). \tag{2.13}$$

2.2.1 Discrete Wavelet Transform

According to (2.9), (2.13), and (2.11), any function $g(t) \in L^2(\mathbb{R})$ can be represented as:

$$\begin{aligned} g(t) &= 2^{j_0/2} \sum_k c_{j0}(k) \varphi(2^{j_0} t - k) + 2^{j/2} \sum_k \sum_{j=j_0}^{\infty} d_j(k) \psi(2^j t - k), \\ g(t) &= \sum_k c_{j_0}(k) \varphi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{\infty} d_j(k) \psi(j, k(t). \end{aligned} \tag{2.14}$$

As in (2.11), the initial scale j_0 is arbitrarily chosen, and defines the lowest signal resolution, containing the coarsest signal information. It is a subspace spanned by the $\varphi_{j_0,k}(t)$. The rest of the $L^2(\mathbb{R})$ is spanned by the wavelet functions, providing the successively finer signal details.

Coefficients $c_{j0,k}$ and $d_{j,k}$ in this wavelet expansion are called the discrete wavelet transform (DWT) of the signal $g(t)$ and they completely describe the $g(t)$. Assuming the wavelet system to be orthogonal, the coefficients can be calculated by inner products:

$$\begin{aligned} c_j(k) &= \langle g(t), \varphi_{j,k}(t) \rangle = \int g(t) \varphi_{j,k}(t) dt, \\ d_j(k) &= \langle g(t), \psi_{j,k}(t) \rangle = \int g(t) \psi_{j,k}(t) dt. \end{aligned} \tag{2.15}$$

2.3 Lifting Scheme Realization of DWT

Wavelet theory has been known for a long time, but widespread application in signal processing was sparked by the work of Mallat, (Mallat, 1988, 1989a,b,c). In collaboration with Meyer, he developed the multiresolution analysis concept for compactly supported wavelets. He successfully linked wavelet and filter bank theory. It was shown that wavelets could be implemented using iterated filter banks. In effect, calculating the DWT came down to a simple task of signal filtering. More details can be found in (Strang and Nguyen, 1996) and (Vetterli and Kovacevic, 1995).

We put our focus on another important step forward in widening wavelet application possibilities. In a series of papers, Sweldens introduced the lifting scheme realization of the wavelet transform, (Stephens, 1976; Sweldens, 1996, 1997; Sweldens and Schroder, 1997). The new approach had several advantages over the conventional filter banks approach. It was an efficient way to calculate the DWT, perfect reconstruction property was guaranteed, all calculations were performed in the spatial domain and it made it easy to introduce adaptivity to a transform. It was shown in (Daubechies and Sweldens, 1998) that any wavelet transform can be decomposed into a finite sequence of simple lifting steps, while in (Kovacevic and Sweldens, 2000) the lifting approach was generalized to arbitrary dimensions.

In the classical approach, all wavelets were translates and dilates of a single mother wavelet function. They are constructed in the frequency domain (Daubechies, 1992) and referred to as the first generation wavelets. In the lifting scheme, wavelet construction is entirely performed in the spatial domain. The benefit is, that it is possible to construct wavelets which are not necessarily translates and dilates of a single mother wavelet function. Such wavelets provide the means to address the problems such as a boundary treatment or processing irregularly sampled signals. Because they inherit all the good properties of the first generation wavelets, while introducing additional features, they are called the second generation wavelets.

The Lifting Scheme

Basic lifting scheme, performing one level of signal decomposition, is shown in Fig. 2.1. It consists of two simple filtering steps – predict (P) and update (U). Purpose of the filter P is to predict values of the odd-indexed samples (Y_o), based on a certain number of even-indexed samples (Y_e). Resulting prediction errors assume the role of the high-pass subband and they are, actually, wavelet coefficients. The coefficients are used in the update step to update properties of the input signal. Filter U is chosen such that the signal average is maintained throughout the decomposition, and output from the "upper" branch represents low-pass subband (Kovacevic and Sweldens, 2000).

Inverse transform, shown in Fig. 2.2, is achieved by simply reversing the order of operations. As it emerges from the lifting scheme architecture, perfect reconstruction

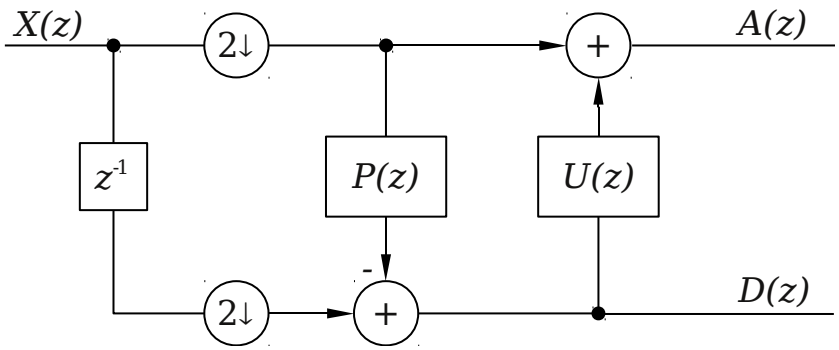


Figure 2.1: Single level lifting scheme decomposition

requirement is guaranteed to be satisfied for any wavelet transform implemented using the concept of lifting.

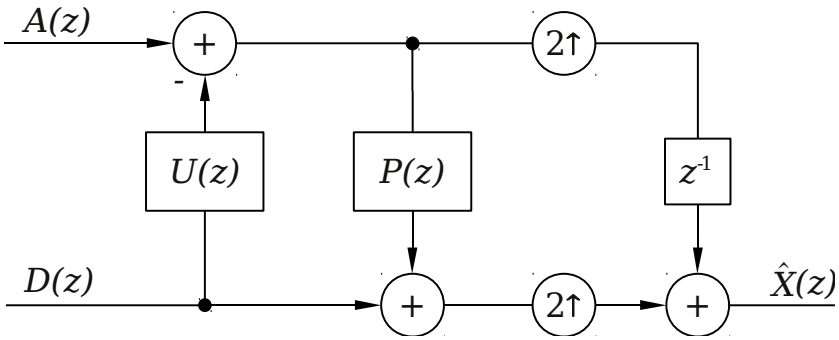


Figure 2.2: Single level lifting scheme reconstruction

To perform full signal decomposition, or reconstruction, iterative filtering is used, the same way as in the classical filter bank realizations of wavelet transforms. The procedure is depicted in Fig. 2.3.

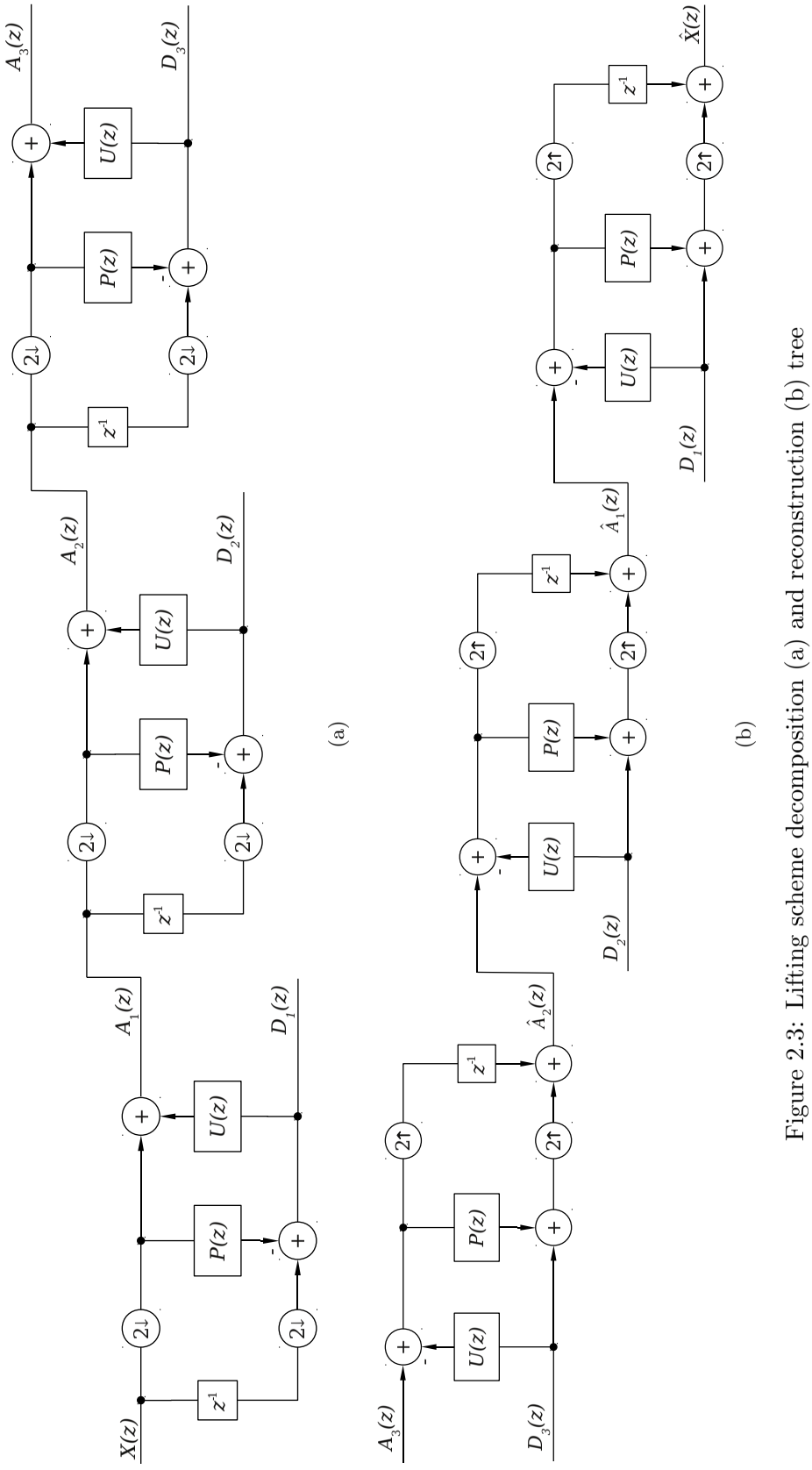


Figure 2.3: Lifting scheme decomposition (a) and reconstruction (b) tree

2.4 Wavelet Transforms in Signal Denoising

Coefficients Thresholding

Application of wavelet transforms in signal denoising exploits two important transform properties. One is the energy packing property – it is to be expected that most of the useful signal information is contained in a relatively small number of larger magnitude wavelet coefficients. The other is the fact that the wavelet transform maps white noise in the signal domain to a white noise in the transform domain. As such, we can assume that the high frequency noise component is almost completely represented by a large number of smaller magnitude wavelet coefficients, concentrated about 0. Should the assumption hold, signal denoising can be performed by altering the values of smaller magnitude coefficients.

Extensive research in the field was carried out by Donoho, (Donoho and Johnstone, 1994, 1995, 1998; Donoho, Johnstone, Kerkyacharian, and Picard, 1995). Wavelet coefficients are altered by introducing a non-linear operation of thresholding. Two commonly used varieties are hard thresholding and soft thresholding, shown in Fig. 2.4. For the hard thresholding, wavelet coefficients lower than the chosen threshold t are set to 0:

$$d_t = \begin{cases} d, & |d| > t \\ 0, & |d| \leq t \end{cases} \tag{2.16}$$

In the soft thresholding case, all coefficients are shrunk by a threshold value t. In effect, all coefficients lower than t in absolute value are set to zero, while the others are pulled toward the origin:

$$d_t = sgn(d)\max(0, |d| - t). \tag{2.17}$$

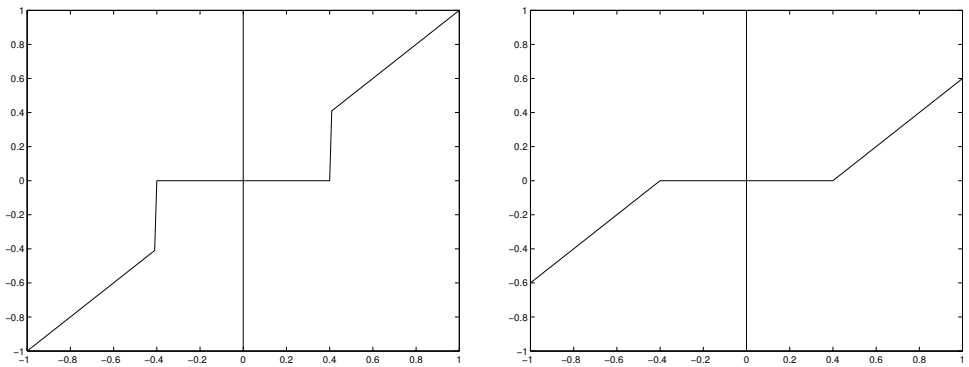


Figure 2.4: Hard (a) and soft (b) thresholding transfer functions

Neither of the methods is found to be superior in every application. Hard thresholding has more respect for sharp changes in a signal, while soft thresholding tends to produce smoother results. For instance, hard thresholding works better for 1-D signals, while soft thresholding performs better for natural images denoising.

Translation Invariant Denoising

As decimated wavelet transforms are not shift-invariant, placement of certain features in a signal can have a huge influence on denoising results. Let us consider the *Haar* wavelet case and an n samples long signal, containing a single step edge between two constant values. If the edge is located near a binary irrational position, such as $n/3$, denoised signal would exhibit significant pseudo-Gibbs phenomena. On the other hand, location of the edge at the $n/2$ position would lead to essentially no pseudo-Gibbs oscillations.

Typically, signals contain much more than a single discontinuity or feature of any other kind, so it is, most often, practically impossible to perfectly align a signal to basis functions. What is an optimal shift for one feature, may represent the worst case scenario for another. Instead of trying to find the perfect shift, transform is calculated across all shifts and results are averaged. The approach was analyzed in (Coifman and Donoho, 1996) and it is an equivalent to a undecimated (stationary) wavelet transform calculation.

Implementation of the undecimated transform using lifting scheme concept is straightforward. Only two modifications to the classical scheme (Fig. 2.1) are needed. We need to remove the decimators, and we have to replace the filters $P(z)$ and $U(z)$, with their upsampled versions: $P(z^2)$ and $U(z^2)$. The implementation is shown in Fig. 2.5. As we progress through the decomposition tree, the filters are further upsampled at each decomposition level.

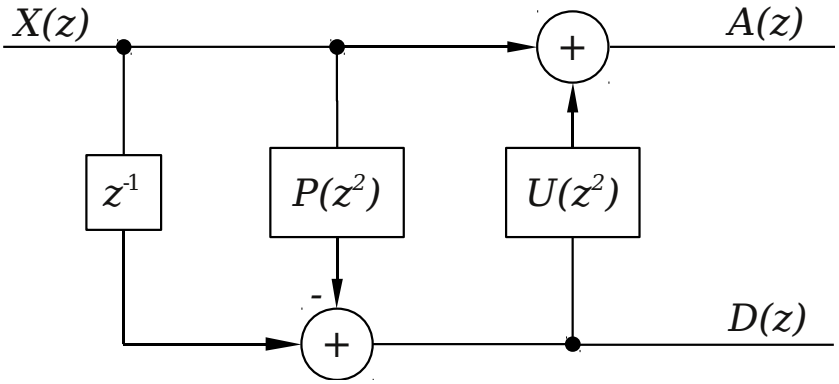


Figure 2.5: Undecimated lifting scheme

Chapter 3

Adaptive Edge Preserving Denoising

In this chapter, we propose a method for signal denoising, based on the adaptive lifting scheme realization of the wavelet transform. The Intersection of Confidence Intervals (ICI) rule, presented by Katkovnik in (Katkovnik, 1999), is explained in details. The ICI rule allows for efficient segregation of signal regions which share similar statistical properties, and it is used as a core of the adaptation algorithm. The main goal of the ICI rule is to prevent usage of lifting filters whose supports span across two or more signal regions of different local properties, as such filters cause undesirable denoising artifacts, especially about edges in a signal.

Before going into details, we will first look at the motivation and basic concepts which lead to the development of the ICI-EPL adaptive denoising scheme.

3.1 Deciding the Right Wavelet

A large number of wavelet basis functions were developed over years. Some important properties of wavelets, from the signal denoising point of view, are support size, smoothness and symmetry. For instance, two different wavelets are shown in Fig. 3.1 – *Haar* and Daubechies 9 (*Db9*). *Haar* wavelet features short support and sharp edges between the two function values. *Db9* has much longer support, it is oscillatory and very smooth. Since reconstructed signal can be viewed as a superposition of coarse signal approximation and scaled and translated wavelet functions, it is easy to predict the effect of using either of the shown wavelets to denoise some typical signals. Scaled and translated *Haar* wavelet very efficiently reconstructs sharp edges but will have difficulties with smooth signal regions. Smoothness and longer support of the *Db9* wavelet represent the opposite case. It efficiently reconstructs the smooth signal regions, but will not be able to do the same with higher frequencies in a signal, eventually leading to a pseudo-Gibbs oscillations at sharp edges. This is illustrated in Fig. 3.2 and Fig. 3.3, in which results of denoising

sample signals with prevailing, respectively, higher and lower frequencies are shown.

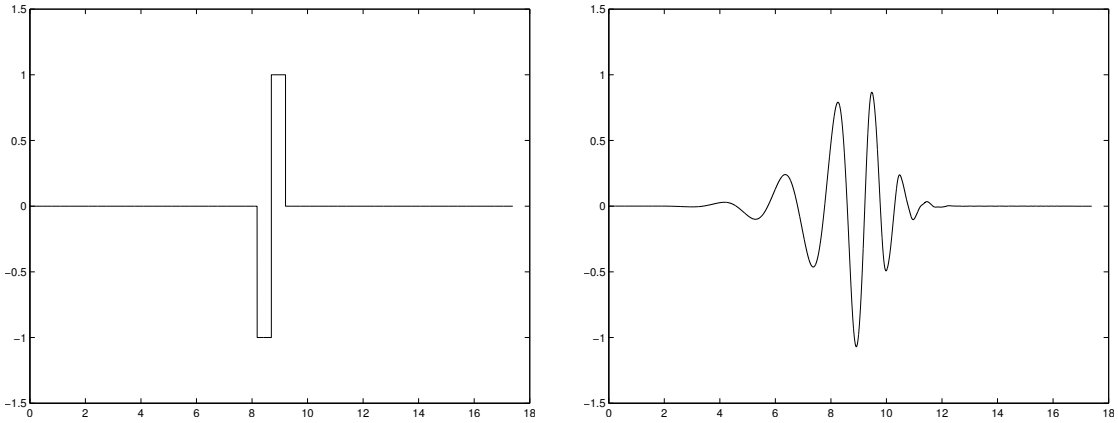


Figure 3.1: Haar (a) and Daubechies 9 (b) wavelet functions

The two examples make general principle for wavelet basis selection obvious. If a given signal is characterized by smooth and slow transitions, i.e. low frequencies, longer and smoother wavelet basis should be chosen. If the signal is characterized by high frequencies, shorter wavelets should be chosen. In reality, the simple principle does not solve the problem of the basis selection, as there are many signal classes beyond the two extreme cases shown in Fig. 3.2 and Fig. 3.3. Given a class of signals, even if the signal properties are same for the whole signal length, there is no simple solution to decide the exact optimal wavelet shape or support length. A more difficult scenario is a class of signals whose properties differ significantly during the course of the signal. In such cases, it would be best to use several different wavelet basis for different signal regions.

It is difficult to address all the concerns without an efficient adaptive algorithm. Lifting scheme concept simplified development of adaptive transforms, and many different approaches to the problem were proposed, as presented in Chapter 1.3.

3.2 Adaptive Lifting Scheme Concept

The adaptive denoising method, we propose, uses the basic principle for a wavelet basis selection – longer wavelets are to be chosen if low frequencies are prevailing in a signal, while shorter wavelets are to be used if higher frequencies are prevailing in a signal. Instead of applying the principle on a signal-by-signal basis, we apply it on a point-by-point basis, in a given signal.

Let us examine the signal in Fig. 3.4. It features four smooth segments, connected by three sharp edges. Our adaptive algorithm should ensure that the longest wavelet which does not span the edge in a signal is used for each signal point. For instance, at edge samples, *Haar* wavelet should be used as it allows for excellent edge reconstruction. As we move away from the edge, longer wavelets should be used to allow for fine reconstruction

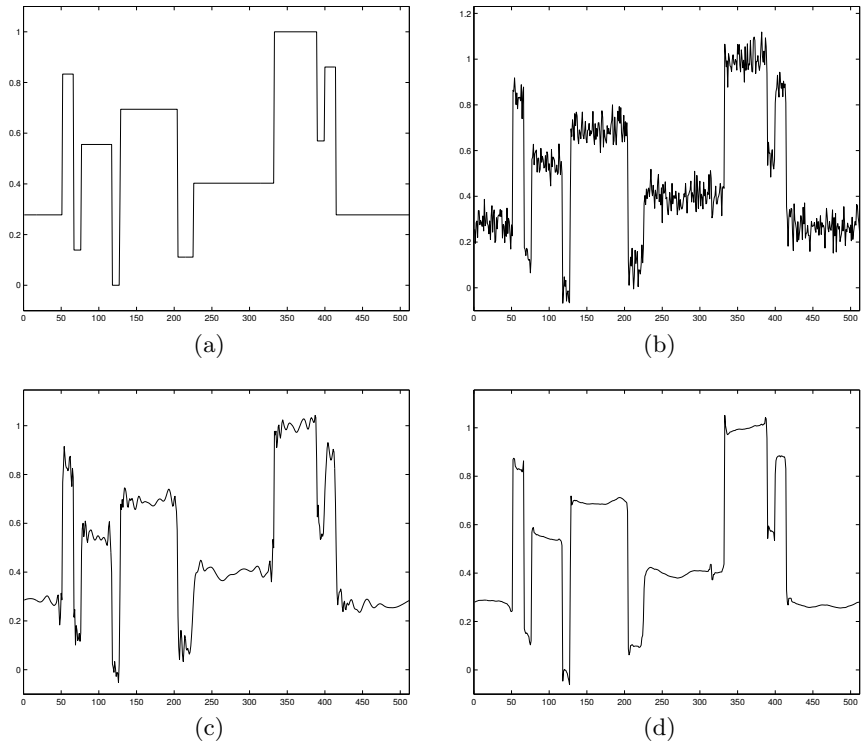


Figure 3.2: Original (a) and noisy (b) *Blocks* signals. Signals denoised using Db9 (c) and Haar (d) wavelets

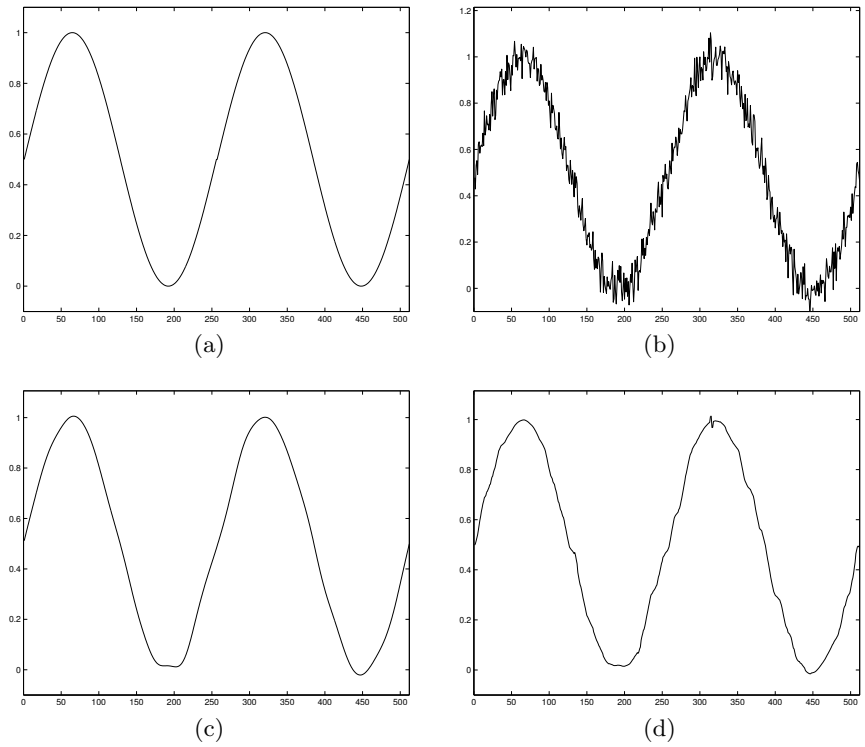


Figure 3.3: Original (a) and noisy (b) Sine signals. Signals denoised using Db9 (c) and Haar (d) wavelets

of smooth signal regions. If done right, the algorithm will bring the good properties of different wavelet basis into a single transform. Denoising scheme based on the transform should perform well on both the smooth regions and edges, contained in the same signal.

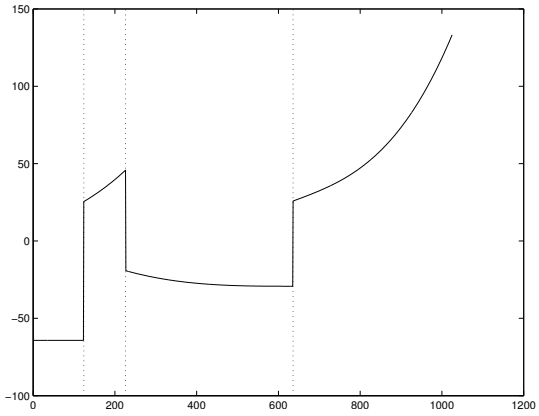


Figure 3.4: Sample signal with variable properties

3.2.1 Definition and Detection of an Edge

Strategy of changing the wavelet basis as we approach the edge, or move away from it, brings about another obstacle – edge detection. Edge detection is a separate research field and there are multiple approaches and solutions to the problem. Before trying to implement any of the detection algorithms, it is necessary to define what does the *edge* exactly mean for our application.

It could be defined as a step edge, a spike or even any high frequency change (drawing another question: What is a high frequency?). We could say that the usable definition of an edge, for this particular application, is that it is the point in a signal which divides two regions of different local properties. It is a very broad definition, which includes much more than what the common edge detectors are trying to detect. To detect the existence of edges conforming to the given definition, we use the statistical method presented in (Katkovnik, 1999) – Intersection of Confidence Intervals (ICI) rule. As explained in more details later in the chapter, the ICI rule help us determine the largest neighborhood about each signal point, which share the same local properties. It is used during the decomposition procedure, for each point independently.

3.2.2 Wavelet Basis Set

Daubechies and Sweldens showed in (Daubechies and Sweldens, 1998) that any wavelet transform with finite filters can be decomposed into a finite sequence of simple lifting steps. Having many lifting steps leads to an unnecessary algorithmic and computational complexity, while adding little benefits. Out of the many existing wavelet families, we decided to use the biorthogonal wavelets which are “native” to the lifting scheme. Only

two lifting steps are required to construct the wavelets: a single predict and a single update step.

Fig. 3.5 shows 4 wavelets used in the adaptive scheme. The wavelets are: *Haar*, *Bior2.2*, *Bior4.4* and *Bior6.6*. Should the adaptive algorithm perform well, the *Haar* wavelet will be used about sharp edges in a signal, while the *Bior6.6* will be used for smoothest signal regions. It is possible to use even longer wavelets but there is no practical advantage of doing so. Even the *Bior6.6* improves denoising performance only for extremely smooth signals.

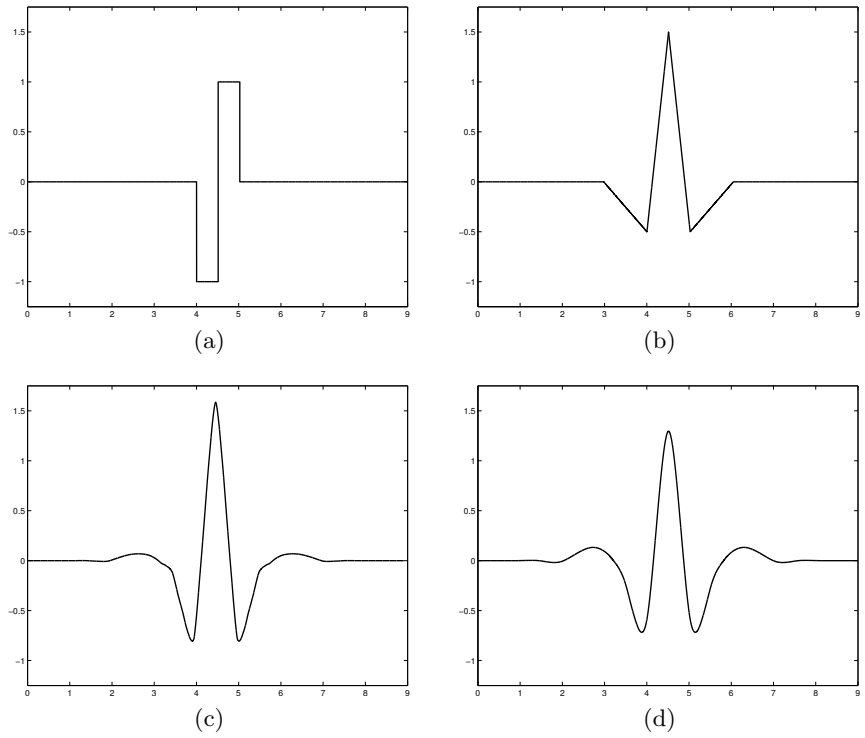


Figure 3.5: Biorthogonal wavelet set for adaptive algorithm: Haar (Bior1.1) (a) Bior2.2 (b) Bior4.4 (c) Bior6.6 (d)

Based on the results of the ICI rule, a set of lifting filters, constructing one of the four wavelets, is chosen and used at each signal point. The adaptation is performed independently for each point and on each scale (decomposition level).

3.2.3 Signal Boundaries

In signal processing, signals are commonly expanded over its boundaries to allow for filtering of the boundary samples with the same filters as the rest of the signal. If wavelet transform was used for processing, it would, essentially, mean using the same wavelet for signal boundary samples as for the midst signal samples. Different types of padding are used for the expansion, each having its own advantages and disadvantages. Since we are constructing the adaptive wavelet transform, featuring second generation wavelets,

there is no need to employ padding. Instead, we can opt to contract the wavelet support as we are getting close to a signal boundary. It is the same effect that can be seen when approaching the edge in a signal. By avoiding signal padding and expansion, the denoising performance in the boundary region can be, depending on a specific signal, greatly improved.

3.3 Intersection of Confidence Intervals – ICI

The ICI is a statistical method which utilizes non-parametric local polynomial approximation (LPA) and confidence intervals (CI) theory to estimate the real value of a sample in a noisy signal. We will only explain part of the method theory which is relevant for the adaptive lifting scheme, while the complete presentation and analysis of the method can be found in (Katkovnik, Egiazarian, and Astola, 2003).

Let the signal:

$$y(x), \quad y \in \mathbb{R}^1, \quad x \in \mathbb{R}^1 \tag{3.1}$$

be a piece-wise polynomial function, sampled on a regular or irregular grid:

$$x = x(n),$$

where index n corresponds to the $n-th$ signal sample. Its noisy observations $z(x)$ are given by:

$$z(x) = y(x) + \epsilon(x), \tag{3.2}$$

where $\epsilon(x) \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is considered to be independent Gaussian white noise.

The aim of the method is to estimate real values of the unknown deterministic $y(x)$, from its noisy observations $z(x)$. Each $y(x)$ is considered to be lying on a polynomial segment. The segment can be approximated by a polynomial of order $m = 0, \ldots, M$. We are only interested in the polynomial value at a given point, so, instead of finding the polynomial parameters, the non-parametric LPA route is taken.

3.3.1 LPA Kernel Calculation

Let the point x be a center point of the LPA, i.e., the point at which function value $y(x)$ is to be estimated using the LPA kernels. Estimate for the point x_s in the neighborhood of x can be looked for as follows:

$$\begin{aligned} &y(x - x_s) = C^T u(x - x_s), \\ &u(x) = (u_1(x), u_2(x), \ldots, u_M(x))^T, \\ &C = (C_1, C_2, \ldots, C_M)^T, \end{aligned} \tag{3.3}$$

where $u(x) \in R^M$ is a vector of linearly independent 1D polynomials, of the power from 0 to M, while $C \in R^M$ is a vector of parameters for this model.

Any set of linearly independent polynomials can be used. For instance, Lagrange polynomials are defined by:

$$L_n(x) = x^n, \tag{3.4}$$

while Hermitian polynomials are defined by:

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}. \tag{3.5}$$

For $m = 3$, the complete set of Lagrange linearly independent polynomials is given by:

$$\begin{aligned} u_1(x) &= 1, \\ u_2(x) &= x, \\ u_3(x) &= x^2, \\ u_4(x) &= x^3, \end{aligned} \tag{3.6}$$

while the complete set of Hermitian linearly independent polynomials is given by:

$$\begin{aligned} u_1(x) &= 1, \\ u_2(x) &= x, \\ u_3(x) &= x^2 - 1, \\ u_4(x) &= x^3 - 3x. \end{aligned} \tag{3.7}$$

Regardless of the polynomial choice, in order to find the coefficients C of (3.3), the following criterion function is applied:

$$J_h(x) = \sum_s w_h(x - x_s)(z(x_s) - y(x - x_s))^2, \tag{3.8}$$

where the window $w_h(x)$:

$$w_h(x) = w(x/h)/h^2 \tag{3.9}$$

formalizes localization of fitting with respect to the center x, while the scale parameter $h > 0$ determines the window size. The window $w(x)$ satisfies the following property:

$$\int w(x)dx = 1. \tag{3.10}$$

Estimates of parameters C are obtained by minimizing the criteria function $J_h(x)$ (3.8) with respect to the C:

$$\hat{C}(x, h) = arg \min_{C \in R^M} J_h(x), \tag{3.11}$$

gives:

$$\begin{aligned}\hat{C}(x,h) &= u_h^{-1}\sum_s w_x(x-x_s)u(x-x_s)z(x_s),\\ u_h &= \sum_s w_h(x-x_s)u(x-x_s)u^T(x-x_s),\\ \hat{C}(x,h) &= (\hat{C}_1(x,h),\ldots,\hat{C}_M(x,h))^T.\end{aligned} \tag{3.12}$$

Having obtained coefficients $\hat{C}$, estimate of the function value at point x_s follows, from (3.3) and (3.11), as:

$$\begin{aligned}\hat{y}_h(x-x_s) &= \hat{C}^T u(x-x_s) = u^T(x-x_s)\hat{C},\\ &= u^T(x-x_s)u_h^{-1}\sum_s (w_h(x-x_s)u(x-x_s)z(x_s)).\end{aligned} \tag{3.13}$$

Let us assume that we orthonormalized the polynomials $u(x-x_s)$ with respect to the window $w_h(x-x_s)$. We shall denote the orthonormalized set of polynomials as $\phi(x-x_s)$. The expression (3.13) can now be written as:

$$\hat{y}_h(x-x_s) = \phi^T(x-x_s)\phi_h^{-1}\sum_s (w_h(x-x_s)\phi(x-x_s)z(x_s)). \tag{3.14}$$

Considering that:

$$\phi_h = I_{MxM}, \tag{3.15}$$

the (3.14) becomes:

$$\hat{y}_h(x-x_s) = \phi^T(x-x_s)\sum_s (w_h(x-x_s)\phi(x-x_s)z(x_s)). \tag{3.16}$$

After substituting:

$$\phi^T(d_s) = \phi^T(x-x_s), \tag{3.17}$$

the function value estimate can be written as:

$$\hat{y}_h(x-x_s) = \sum_s (w_h(x-x_s)\phi^T(d_s)\phi(x-x_s)z(x_s)). \tag{3.18}$$

The notation of (3.18) can be further simplified to:

$$\begin{aligned}\hat{y}_h(x-x_s) &= \sum_s g_h(x-x_s)z(x_s),\\ g_h(x-x_s) &= w_h(x-x_s)\phi^T(d_s)\phi(x-x_s),\end{aligned} \tag{3.19}$$

showing that the estimate can be obtained by filtering the noisy signal with the finite

impulse response filter, given by its impulse response $g_h(x - x_s)$.

Orthonormalization of the Polynomials

It was assumed in (3.15) that we orthonormalized the set of linearly independent polynomials $u(x - x_s)$ such that they are orthonormal with the window $w_h(x - x_s)$. The orthonormalization was accomplished using the Gram-Schmidt process, presented below.

Given a set of linearly independent functions $\{u_n\}_{n=0}^{\infty}$, let $\{\psi_n\}_{n=0}^{\infty}$ denote the orthogonalized functions and $\{\phi_n\}_{n=0}^{\infty}$ denote the orthonormalized functions. Let us further define that:

$$\begin{aligned} \psi_0(x) &\equiv u_0(x), \\ \phi_0(x) &\equiv \frac{\psi_0(x)}{\sqrt{\int \psi_0^2(x) w(x) dx}}. \end{aligned} \tag{3.20}$$

Now take:

$$\psi_1(x) = u_1(x) + a_{10}\phi_0(x), \tag{3.21}$$

where it is required that:

$$\int \psi_1 \phi_0 w dx = \int u_1 \phi_0 w dx + a_{10} \int \phi_0^2 w dx = 0. \tag{3.22}$$

By definition,

$$\int \phi_0^2 w dx = 1, \tag{3.23}$$

so, from (3.21) and (3.22), follows:

$$a_{10} = -\int u_1 \phi_0 w dx. \tag{3.24}$$

The first orthogonalized function can now be found as:

$$\psi_1 = u_1 - \left[\int u_1 \phi_0 w dx\right] \phi_0, \tag{3.25}$$

while the first orthonormalized function is given by:

$$\phi_1 = \frac{\psi_1}{\sqrt{\int \psi_1^2 w dx}}. \tag{3.26}$$

All the other orthogonal functions can be found as:

$$\begin{aligned} \psi_i &= u_i + a_{i0}\phi_0 + a_{i1}\phi_1 + \cdots + a_{i(i-1)}\phi_{i-1}, \\ a_{ij} &\equiv -\int u_i \phi_j w dx, \end{aligned} \tag{3.27}$$

while the corresponding normalized functions are given by:

$$\phi_i = \frac{\psi_i}{\sqrt{\int \psi_i^2 w dx}}. \tag{3.28}$$

Sample LPA Kernel

To illustrate the procedure for the LPA kernel generation, we consider an example in which the true value of a point is estimated using symmetrical rectangular window, $h = 5$ signal points wide, and Lagrange polynomials of order $m = 0$ to 2.

The starting point for kernel calculation is the expression (3.19):

$$\begin{aligned} \hat{y}_h(x - x_s) &= \sum_s g_h(x - x_s) z(x_s), \\ g_h(x - x_s) &= w_h(x - x_s)\phi^T(d_s)\phi(x - x_s). \end{aligned} \tag{3.29}$$

Let us assume that the center point of the LPA is $x = x_s$, so that $d_s = 0$. If we, also, make a substitution: $l = x - x_s$ in (3.19), we get:

$$g_h(l) = w_h(l)\phi^T(0)\phi(l). \tag{3.30}$$

Symmetrical window function is given by:

$$w_h(l) = \frac{1}{2L+1}, \quad l \le |L|, \quad L = \frac{h-1}{2}, \tag{3.31}$$

while a complete set of linearly independent Lagrange polynomials orthonormal with the $w_h(l)$ (3.31) is:

$$\begin{aligned} \phi_1(l) &= 1, \\ \phi_2(l) &= \sqrt{\frac{3}{L(L+1)}} l, \\ \phi_3(l) &= \sqrt{\frac{5}{L(L+1)(2L+3)(2L-1)}} (3L^2 - L(L+1)). \end{aligned} \tag{3.32}$$

The final expression for the LPA kernel coefficients is then:

$$g_h(l) = (1 + 5\frac{L(L+1) - 3l^2}{(2L+3)(2L-1)})\frac{1}{2L+1}, \quad l \le |L|, \tag{3.33}$$

And the LPA kernel for our example is:

$$g_h(l) = [-0.085 \quad 0.343 \quad 0.486 \quad 0.343 \quad -0.085]. \tag{3.34}$$

3.3.2 Idea of the ICI

Let us consider the LPA kernels from the example above. Symmetrical window with $|L| = 2$ and estimation order $m = 2$ are used. In case the center point of LPA lies on a polynomial segment whose order is $\leq m$, the kernels will provide perfect estimate of the function real value, for any window size $h > m$. For instance, given a point on a parabola, to get $\hat{y}(x) = y(x)$, it is enough to have $h = 3$. In this case, $L = 1$ would suffice.

In reality, the original signal $y(x)$ is composed of many different polynomials. Also, it is corrupted by noise and we are only given its noisy observations $z(x)$. In such cases, it would be difficult to expect that the polynomials contained in $y(x)$ could be accurately reconstructed based on the $m + 1$ points. In fact, it would take $h = \infty$ for the perfect signal reconstruction to be possible. Although the perfect reconstruction is not likely scenario, widening of the estimation window is a rewarding approach for improving the $\hat{y}(x)$ accuracy. Additional signal points in the wider window allow for better polynomial approximation and, from the statistical point of view, lower estimate variance.

As long as all the window points lie on the same polynomial segment, of order $\leq m$, the LPA kernels are regarded to be unbiased estimators. Lower variance is then, indeed, reflected as a higher probability that the function value estimate is closer to the real value $y(x)$.

Influence of the estimation bias is another factor to consider. The problem with growing estimation window arises when the enlarged window includes points which lie on a different polynomial segment. LPA kernel estimators become biased and variance alone cannot be looked at as a measure of accuracy. Should the window be further extended, estimator bias could cause significant worsening of the $\hat{y}(x)$, regardless of the low estimate variance. This is depicted in Fig. 3.6.

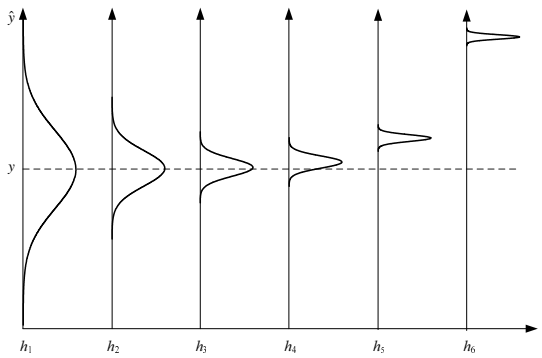


Figure 3.6: Probability distribution functions of $\hat{y}(x)$, for growing window sizes h_i

It is evident that an unbiased estimator can result in an unacceptably high estimate

variance, while at the same time lowering of the variance, by extending the estimation window, can result in an unacceptable estimation bias. The ICI rule considers both of the variables and, as shown in (Stankovic, 2004), provides optimal bias-to-variance trade-off.

Determining Support Size – ICI Rule

Since the LPA kernels (3.19) are, basically, finite impulse response (FIR) filters, the estimation window can also be viewed as a filter support. From now on, the term "support " will be used in place of the estimation window.

Let us examine a case in which the kernels are used to estimate the real function value, $y(x)$. Associated with the resulting estimates are their upper (U) and lower (L) confidence interval (CI) limits, given by:

$$\begin{aligned} U &= \hat{y}(x) + z_c \cdot \sigma_{\hat{y}}, \\ L &= \hat{y}(x) - z_c \cdot \sigma_{\hat{y}}, \end{aligned} \tag{3.35}$$

where the critical level $z_c = \chi_{1-\alpha/2}$ is the $(1-\alpha/2)-th$ quantile of the standard Gaussian distribution and defines the probability that the true value $y(x)$ is contained within the CI limits.

To make the expression (3.35) more accurate, the bias has to be considered. (Stankovic, 2004) showed that the upper bound of the estimation bias is given by:

$$\begin{aligned} \bar{\omega}(x,h) &< \gamma \cdot \sigma_{\hat{y}}, \\ \gamma &= \left(\sqrt{(m+1)}\,\right)^{-1}. \end{aligned} \tag{3.36}$$

Then, from (3.35) and (3.36), the expressions for CI limits are:

$$\begin{aligned} U &= \hat{y}(x) + \Gamma \cdot \sigma_{\hat{y}}, \\ L &= \hat{y}(x) - \Gamma \cdot \sigma_{\hat{y}}, \end{aligned} \tag{3.37}$$

where $\Gamma = \gamma + \chi_{1-\alpha/2}$ is a free parameter, which defines the method sensitivity. Too large or too small Γ parameter value, respectively, results in oversmoothing or undersmoothing the signal. The value is seen as a compromise, which has heavy influence on the final results quality.

The ICI algorithm demands that we calculate CIs for estimators of growing supports:

$$\mathbb{H} = \{H_i,\ i = 1, \ldots, N\},$$

with length h_i of each succeeding support being larger than the previous one:

$$\{h_i \mid h_i < h_{i+1}\}.$$

Relevant part for the ICI method are the CI extreme values. Maximum lower CI limit

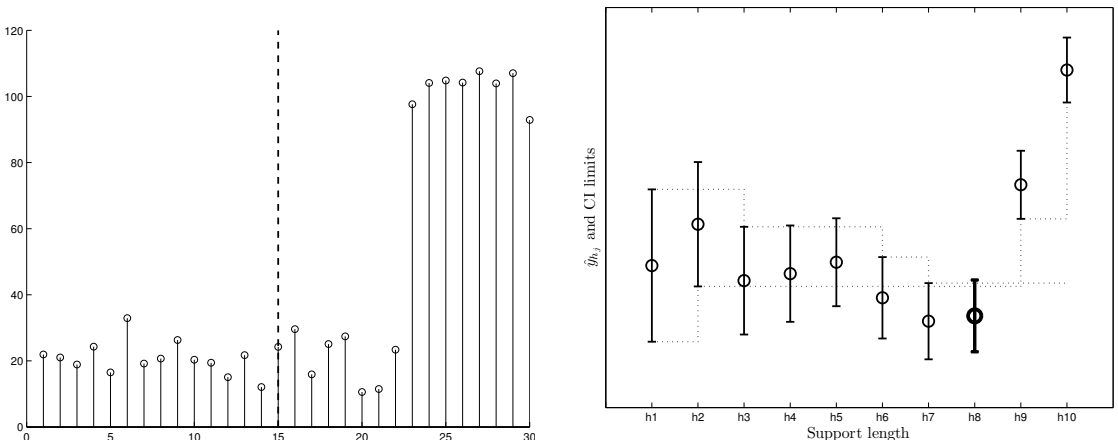
and minimum upper CI limit values are recorded for each H_i (Katkovnik, Egiazarian, and Astola, 2002):

$$\begin{aligned} \bar{L}_{i+1} &= max[\bar{L}_i, L_{i+1}], \\ \underline{U}_{i+1} &= min[\underline{U}_i, U_{i+1}], \\ i = 1, 2, \ldots, N,\ \bar{L}_1 &= L_1,\ \underline{U}_1 = U_1. \end{aligned}$$

The ICI rule states that there is a high probability that the estimator producing the most accurate estimate $\hat{y}(x)$ is the one with the largest support H_i, for which the condition:

$$\bar{L}_i \leq \underline{U}_i \tag{3.38}$$

is still satisfied. The respective support is denoted as H^+. For the next larger filter, there would be no more intersection of $\bar{L}_i$ and $\underline{U}_i$ and it is to be expected that the additional samples added to the H^+ belonged to a signal region with different local properties. This is better illustrated in Fig. 3.7. Fig. 3.7 (a) shows part of a noise corrupted two-value signal, for which the ICI rule is used to estimate the real value of the sample at $x = 15$ (as denoted by the vertical dashed line). Fig. 3.7 (b) shows confidence interval bars in case the support is grown in the right direction only. The support H_9 is the first one to include the sample from the higher signal level. In the CI plot, it can be identified as a breach of the condition (3.38). As per the ICI rule, the filter with support H_8 is to be chosen as the estimator for $y(15)$, as it is the filter with the largest support for which the condition (3.38) still holds.



(a) (b)

Figure 3.7: Sample noisy signal (a) and CI limits for $\hat{y}(15)$ (b). Filter support is growing to the right

The ICI rule should ensure that the chosen filter support, generally, does not span across the edges between the polynomial segments in a signal. In case it does, number of points from the neighboring segments included in the support should stay very low, and

their influence too small to significantly impact the estimation bias.

3.4 Edge Preserving Lifting Scheme – ICI-EPL

In Chapter 3.1 and 3.2 we explained the rationale and the goals for the ICI-EPL adaptive denoising algorithm. We will now make an implementation of the algorithm using the ICI rule as the core of the adaptive algorithm.

Let us, again, consider the lifting scheme realization of the wavelet transform (Fig. 2.1). In the predict step, filter P is used to predict values of the odd-indexed signal sample, based on a certain number of even-indexed samples, and vice versa. Because, for conventional wavelets, filter P is same for every signal sample, we get a performance trade-off – longer filter P results in better performance in smooth signal regions, while shorter P results in better performance about edges in a signal. In order to avoid the trade-off, the ICI rule can be used to determine the filter P support for each signal sample, independently.

3.4.1 Lifting Filters

As explained in 3.3.2, the ICI rule considers LPA kernels (3.19), while lifting filters used in the lifting scheme are Neville filters. There are two options – either Neville filters will be used for the ICI rule, or the (3.19) LPA type kernels will be used in the lifting scheme. Let us briefly examine both options.

Neville Filters and the ICI Rule

Essentially, Neville filters of order N are also LPA kernels. They are interpolating filters which are able to perfectly reconstruct any polynomial whose order is $m < N$. The same is true for the LPA kernels. The difference between the ICI utilization of LPA kernels and Neville filters is in the relationship of order of estimation and support size.

LPA kernels used for the ICI rule all share the same estimation order, set to a value of $m = const$, regardless of the kernel support size. For the support size of $h = m + 1$, the kernels are exactly the same as the Neville filters of order $N = m + 1$. As the support of Neville filters grow, so does their order of estimation, which is always equal to: $m = N - 1$. If used for the ICI rule, both filter types must have growing support, but in the LPA kernel case, estimation order is fixed, while for the Neville filters it grows with the support. Important consequence is relationship of estimate variance to support size. Both estimator types are finite impulse response (FIR) filters and the respective estimate variance is derived as follows:

$$\begin{aligned} \hat{y}(x) &= \sum_k g_h(k) y(x-k), \\ Var(\hat{y}) = \sigma^2_{\hat{y}} &= \sum_k g^2_h(k) \cdot \sigma^2_{y(x-k)}. \end{aligned} \tag{3.39}$$

Considering that:

$$\sigma_{y(x-k)} = \sigma_\epsilon,$$

the variance becomes:

$$Var(\hat{y}) = \sum_k g_h^2(k) \cdot \sigma_\epsilon^2,$$

and finally:

$$\begin{aligned} Var(\hat{y}) = \sigma_{\hat{y}}^2 &= \sigma_\epsilon^2 \cdot \sum_k g_h^2(k) \\ \sigma_{\hat{y}} &= \sigma_\epsilon \cdot \sqrt{\sum_k g_h^2(k)}. \end{aligned} \tag{3.40}$$

The variance is directly proportional to the sum of squared filter/kernel coefficients, and, as per (3.35), the confidence interval width is directly proportional to the variance. In the LPA kernel case the variance and the CI width are getting lower as the support grows, while for the Neville filters, they are getting higher. The ICI rule efficiency is based on the shrinking CI widths, as shown in Fig. 3.7. The effect of expanding CI widths can be observed in Fig. 3.8. The Figure shows plot of CI bars for a sample signal, when using Neville filters of order $N \in 1, 2, 3, 4, 5, 6, 7$. It can be seen that the expanding CIs are detrimental to the efficiency of the ICI rule, as it slows or even disables the breaching of the condition (3.38). We can conclude that the Neville filters are not suitable for the ICI rule utilization.

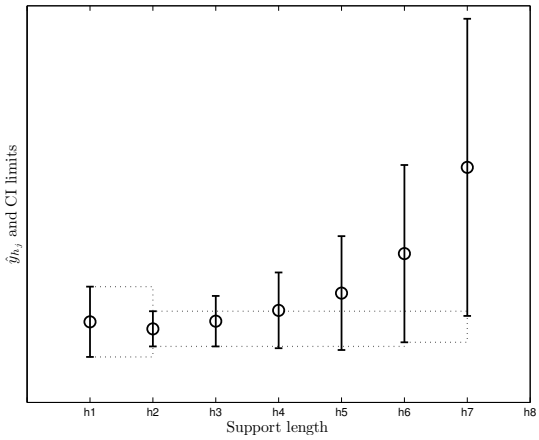


Figure 3.8: Effect of growing variance of Neville filters to the ICI rule efficiency

LPA Kernels and the Lifting Scheme

LPA kernels as proposed in (Katkovnik, 1999) cannot be directly used in the lifting scheme. Katkovnik assumes that the kernel uses every support point for the approximation. Also, he assumes that, in (3.19), the center point of the LPA is always equal to x_s. It leads to

$d_s = 0$ and modified expression (3.19):

$$g_h(x - x_s) = w_h(x - x_s)\phi^T(0)\phi(x - x_s). \tag{3.41}$$

Considering the symmetrical estimation window H of length $h = 5$, the effect of the assumptions is that the estimators are used to estimate the value of signal sample $H(3)$ based on the values of all points of H, including $H(3)$.

Filter P in the lifting scheme does not have all the samples at its disposal. It can only use samples from one signal phase, to predict value of one of the samples from the other signal phase. To achieve this, only minor additions to the Katkovnik approach are necessary.

Let us examine the case if the rectangular asymmetrical window function:

$$w_h(x - x_s) = 1/h,$$

as in Fig. 3.9, is assumed. Points in the support H all belong to the same signal phase. Samples from the other signal phase lie at the half-integer grid points. The kernels (3.41) must be modified in order to estimate value of one of the mid-points. It can be easily accomplished by proper selection of the d_s in (3.19). It shall not be set to 0 but to a half-integer value, defining the point in the other signal phase, whose value is to be estimated. For instance, given a support length $h = 4$, and setting the $d_s = 1.5$ results in a symmetrical kernel, estimating the point in the center of support. Upsampled kernels can now be used in place of the filter P.

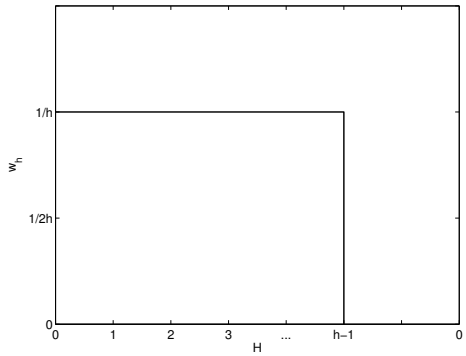


Figure 3.9: Asymmetrical rectangular window

Inclusion of the LPA kernels into the lifting scheme has obvious advantage of allowing for the ICI rule to be efficiently used to determine the filter P support, which can now be calculated separately for each signal sample. However, there is another characteristic of the proposed denoising algorithm, which limits the potential of the approach. Obtaining the best possible prediction in the predict step of the lifting scheme is not enough to achieve excellent denoising performance. Purpose of the filter P is not only the prediction but also, and in the first place, wavelet construction. As we already pointed out, wavelet

basis choice has a major influence on denoising performance, and we must examine the wavelets which get constructed with the LPA kernels.

Fig. 3.10 shows wavelets in case LPA kernels with estimation order $m = 1$ are used. Plots are provided for support lengths $h \in \{4, 6, 8\}$. For $h = 2$, the LPA kernel is identical to the Neville filter, and so are the resulting wavelets. Fig. 3.11 shows the case for $m = 2$ and support lengths $h \in \{6, 8\}$. Again, for $h = 4$ the LPA kernel and Neville filters are identical. If we compare the resulting wavelets with the ones from Neville filters, shown in Fig. 3.5, we can see some general similarity in the shape but there is a significant difference in regularity. Wavelets constructed using Neville filters are very smooth, while the wavelets from LPA kernels are not nearly as smooth and have numerous singularities.

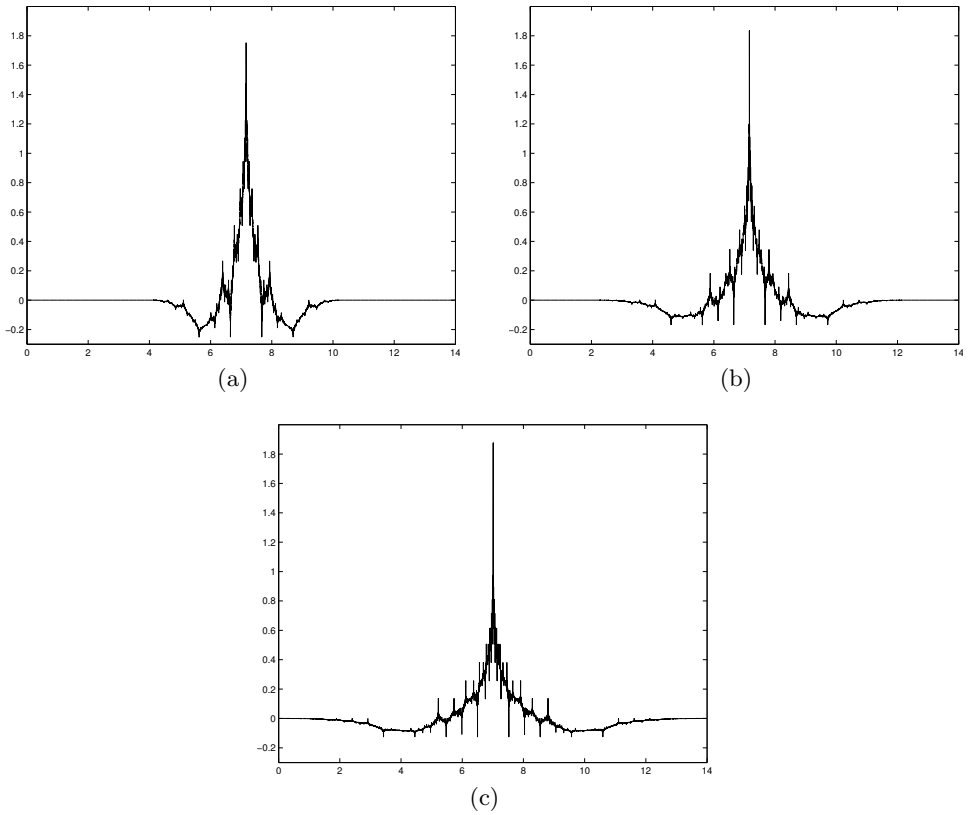
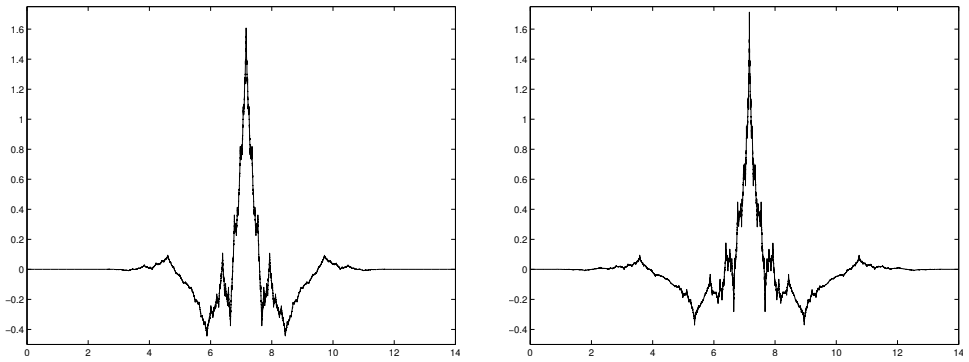


Figure 3.10: Synthesis wavelets from LPA kernels with estimation order $m = 1$. Support length: $h = 4$ (a), $h = 6$ (b) and $h = 8$ (c)

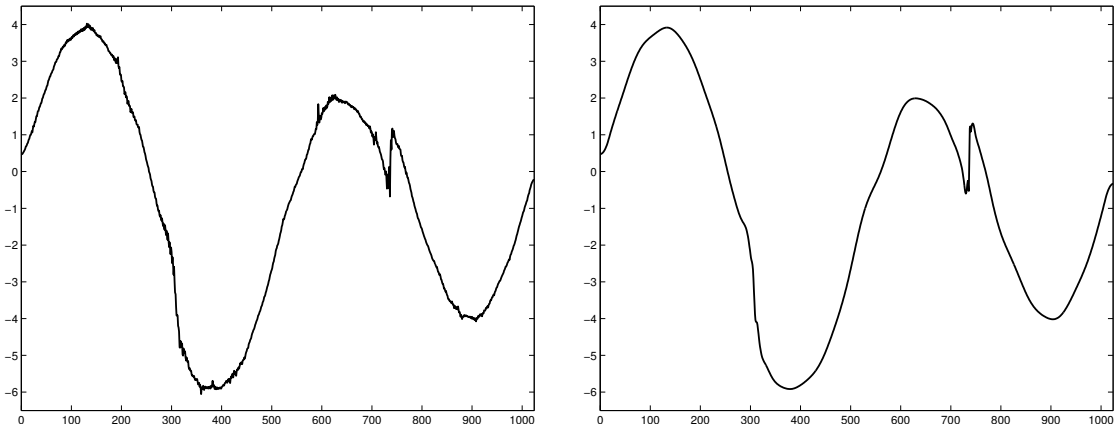
Given that the reconstructed signal is a sum of the coarse signal approximation and translated and scaled wavelet functions, we can anticipate the negative influence of such wavelets on general denoising performance. It becomes nearly impossible to reconstruct the smooth signal regions without the denoising artifacts in form of, mostly, small fluctuations about the real signal value, being clearly visible. The effect is shown in Fig. 3.12. The example compares two denoised *HeaviSine* signals. One is denoised using the wavelets resulting from Neville filters with support length $h = 8$, while the other is denoised using the wavelets resulting from LPA kernels with $m = 2$ and the same support



(a) (b)

Figure 3.11: Synthesis wavelets from LPA kernels with estimation order $m = 2$. Support length: $h = 6$ (a) and $h = 8$ (b)

length, $h = 8$.



(a) (b)

Figure 3.12: *HeaviSine* signal denoised using wavelet transforms based on the (a) LPA kernel, $m = 2$, $h = 8$ and (b) Neville filter $h = 8$

Combining the Advantages

We showed that the Neville filters work excellent for wavelet construction. Wavelets with smaller support feature sharp edges and are good for reconstructing higher frequencies in a signal, while wavelets with larger supports are very smooth and can easily reconstruct lower frequencies in a signal. Unfortunately, Neville filters are not suitable to be used for the ICI rule, because estimate variance rises with the filter support. On the other hand, LPA kernels can be used with the ICI rule to achieve very efficient support size determination, but wavelets they construct are inferior to the ones constructed by Neville filters. They have many singularities, leading to visible artifacts in the denoised signal. The logical solution would be to try to explore the advantages of both filter types. To

achieve this, we propose the signal decomposition in the lifting scheme to be broken into two separate steps.

In the first step, on each scale and for each point, the ICI rule with the symmetrical LPA estimators of order m, with set of growing supports:

$$\mathbb{H}_r = \{H_{m,i},\ i = 1, \ldots, N\}$$

of length:

$$\{h_{m,i} \mid h_{m,i} < h_{m,i+1}\}$$

are used to find the largest admissible support (denoted by '+') H_m^+ of length h_m^+. Based on the H_m^+ even-indexed points (denoted as $H_{m,e}^+$), the estimators predict value of one of the H_m^+ odd-indexed points (denoted as $H_{m,o}^+$), and vice versa. Since symmetrical estimators are assumed, the supports $H_{m,i}$ also grow symmetrically. Ideally, samples which comprise the support H_m^+ would not belong to two (or more) signal regions with different local properties.

In the second step, actual decomposition is performed by substituting the chosen LPA kernels with Neville filters of the same support H_m^+. The substitution can be justified by the fact that the resulting smooth wavelets will have at least the same number of vanishing moments as if the LPA kernel estimators were used. Possible additional vanishing moments would provide smoother reconstruction but would have no negative impact on edge preservation.

Choosing the Estimation Order

Basically, any estimation order m can be chosen for the LPA kernels, although differences in actual denoising performance are not very significant. Another factor to consider when choosing the estimation order is a potential for support growth. Because employing longer wavelets do not bring further performance improvements, we have decided the longest wavelet to be the *Bior6.6*. In effect, the longest filter support that can be used for the ICI rule is of length $h = 6$. Too large estimation order m would, thus, limit the potential for support growth. For instance, should we chose the order $m = 3$, minimum support length would have been $h = m + 1 = 4$. Since the support is grown symmetrically, only two supports would have been considered by the ICI algorithm, which is not enough to explore its benefits or ensure reasonable reliability.

As well as limiting support growth potential, larger estimation order values also entirely exclude smaller wavelets. It means that longer wavelets would be used even at a step edge. This is in harsh contradiction with the basic idea of adaptation and can severely lower transform performance about edges in a signal.

The simulations showed that the estimation order $m = 1$ proved to be enough for optimal denoising performance and it is the order we propose to be used in the ICI-EPL.

Support Verification

Since we opted to set the kernel estimation order to $m = 1$, the resulting symmetrical estimators are simple averaging filters:

$$P_N(z) = \frac{1}{N} \sum_{k=0}^{N-1} z^{-k}.$$

As a result, the more the support grows, the lesser is the impact of added samples ($H_{m,i+1} \setminus H_{m,i}$) on the prediction result. It means that some edges might still show up in the final support H_m^+. Small denoising performance improvements can be achieved by further examining the H_m^+.

In order to verify the support and increase probability that it does not contain any sudden changes in signal statistics, two asymmetrical LPA kernel estimator sets are introduced. Each set is independently used in the ICI algorithm as a set of filters with growing supports. They predict, respectively, values of the second and next to the last points of H_m^+. Their supports are growing from boundaries of the H_m^+ to the right and to the left, respectively. The ICI algorithm results in supports H_m^{L+} and H_m^{R+}, whose lengths (h_m^{L+} and h_m^{R+}) are $\leq h_m^+$. The asymmetrical estimators ensure that samples added in each step of the ICI algorithm carry higher weight, which increases the probability of edge recognition.

In case that:

$$\min[h_m^{L+}, h_m^{R+}] < h_m^+,$$

it is possible that the ICI rule did not indicate change in signal properties fast enough and that the H_m^+ contains signal samples from different segments. Our strategy is to contract the H_m^+ by one filter tap from both sides. The verification step is then repeated until:

$$h_m^{L+} = h_m^{R+} = h_m^+$$

.

3.5 Results and Discussion

To investigate performance of the proposed ICI-EPL denoising scheme, we use 6 different test signals: *Blocks*, *Bumps*, *Doppler*, *HeaviSine*, *Piece-Polynomial* and *Piece-Regular*. The signals are shown in Fig. 3.13. They are generated as if they were samples of the respective continuous time functions, taken at different sampling frequency. We chose 5 different frequencies to obtain discrete time test signals, whose lengths are from the following set: $\{256, 512, 1024, 2048, 4096\}$.

Additive white Gaussian noise was added to each of the signals. Noise was generated at 4 levels, equaling $\sigma_\epsilon \in \{5\%, 10\%, 15\%, 20\%\}$ of total signal magnitude. Illustration

of actual signal corruption caused by noise is shown in Fig. 3.14. The figure shows four plots of the *HeaviSine* signal, containing white Gaussian noise at levels from the above-mentioned set.

For comparison, signals are denoised using the original ICI method, as proposed in (Katkovnik, 1999), the ICI-EPL and 5 well known conventional wavelet transforms: *Haar*, *Bior2.2*, *Bior4.4*, *Db3* and *Db9*. *Haar* and *Bior* wavelets were shown in Fig. 3.5, while *Db9* was shown in Fig. 3.1. Shape of the *Db3* wavelet is depicted in Fig. 3.15.

Proper selection of the parameter Γ value is essential for achieving high efficiency of the proposed ICI-EPL algorithm. In this performance analysis, signals were independently denoised for each of the Γ values

$$\Gamma \in [0.5, 5],$$

in increments of 0.1. The best results were chosen and presented. Automated selection of the parameter Γ value will be discussed in details in the next chapter.

To eliminate influence of threshold selection on the transform efficiency performance comparison, all the wavelet transforms use the optimal threshold value. It is chosen by using thresholds from the $[\sigma_\epsilon, 10\sigma_\epsilon]$ range and selecting the best case, separately, for each of the transforms.

We will thoroughly examine the denoising scheme performance for the signals whose length is 1024 points. Two levels of additive white Gaussian noise will be considered: $\sigma_\epsilon \in \{5\%, 10\%\}$ of total signal magnitude. Objective performance measures for the examined cases, in terms of the root mean square error (RMSE), are presented in Tab. 3.5. Results for all other signal length and noise level combinations are given in Appendix A.

Table 3.1: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon \in \{5\%, 20\%\}$ of total signal magnitude. Signal length is 1024 points. RMSE multiplied by a 10^2

	$\sigma_\epsilon = 5\%$		$\sigma_\epsilon = 20\%$	
	Optimal wavelet	ICI-EPL	Optimal wavelet	ICI-EPL
Blocks	1.22	**0.89**	5.42	**5.21**
Bumps	**1.91**	1.94	**6.57**	6.82
Doppler	**1.66**	1.69	5.59	**5.51**
HeaviSine	1.14	**1.09**	**3.24**	3.61
Piece-Polynomial	1.21	**0.95**	4.86	**4.83**
Piece-Regular	1.72	**1.41**	**5.72**	5.78

Blocks signal

First signal to be analyzed is the *Blocks* signal. It is a piece-wise constant signal, characterized by sharp step edges between constant-level signal regions. *Haar* wavelet transform is optimal transform for denoising such signals. It has the most compact

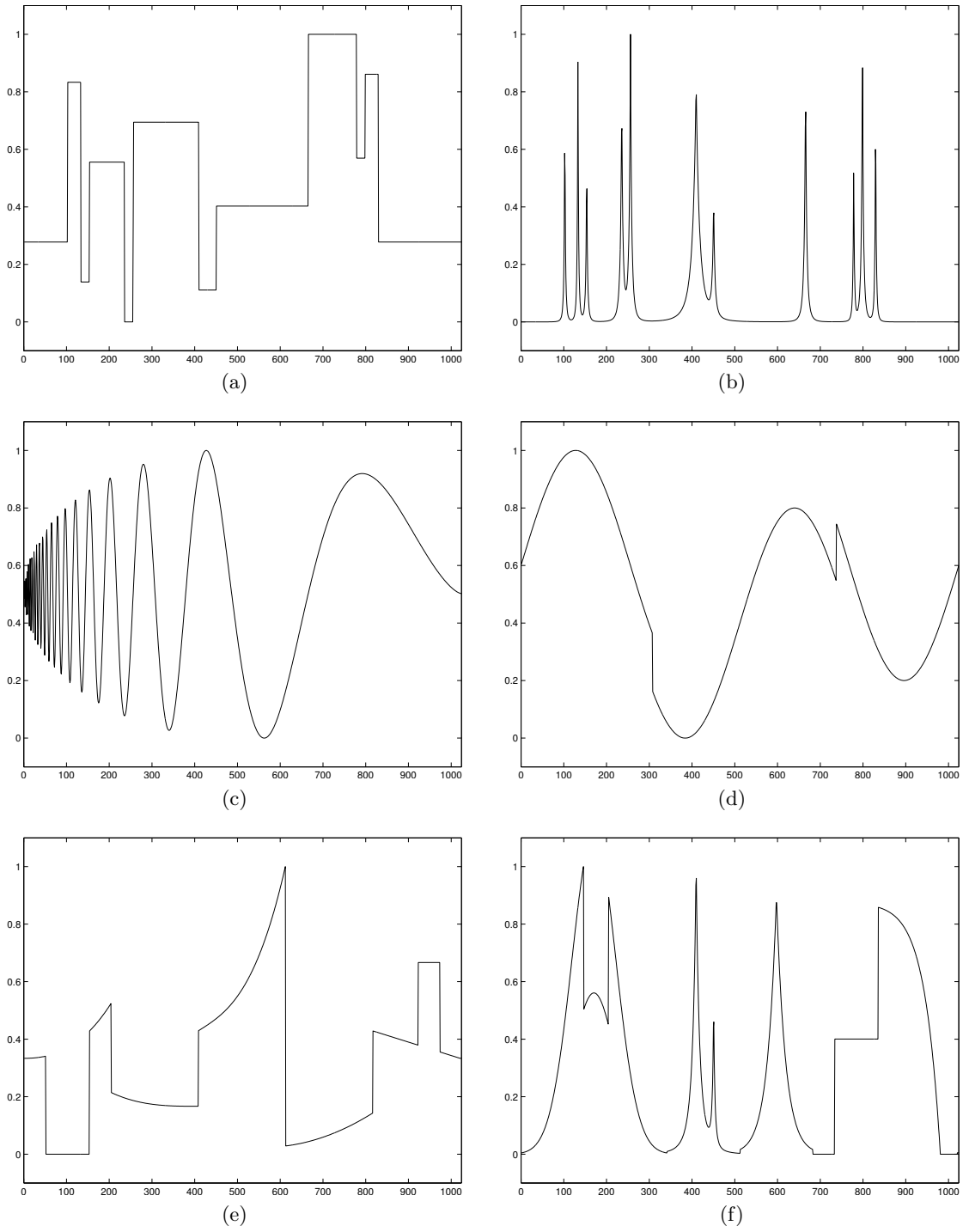


Figure 3.13: Test signals for evaluation of denoising performance: (a) *Blocks*, (b) *Bumps*, (c) *Doppler*, (d) *HeaviSine*, (e) *Piece-Polynomial* and (f) Piece-Regular

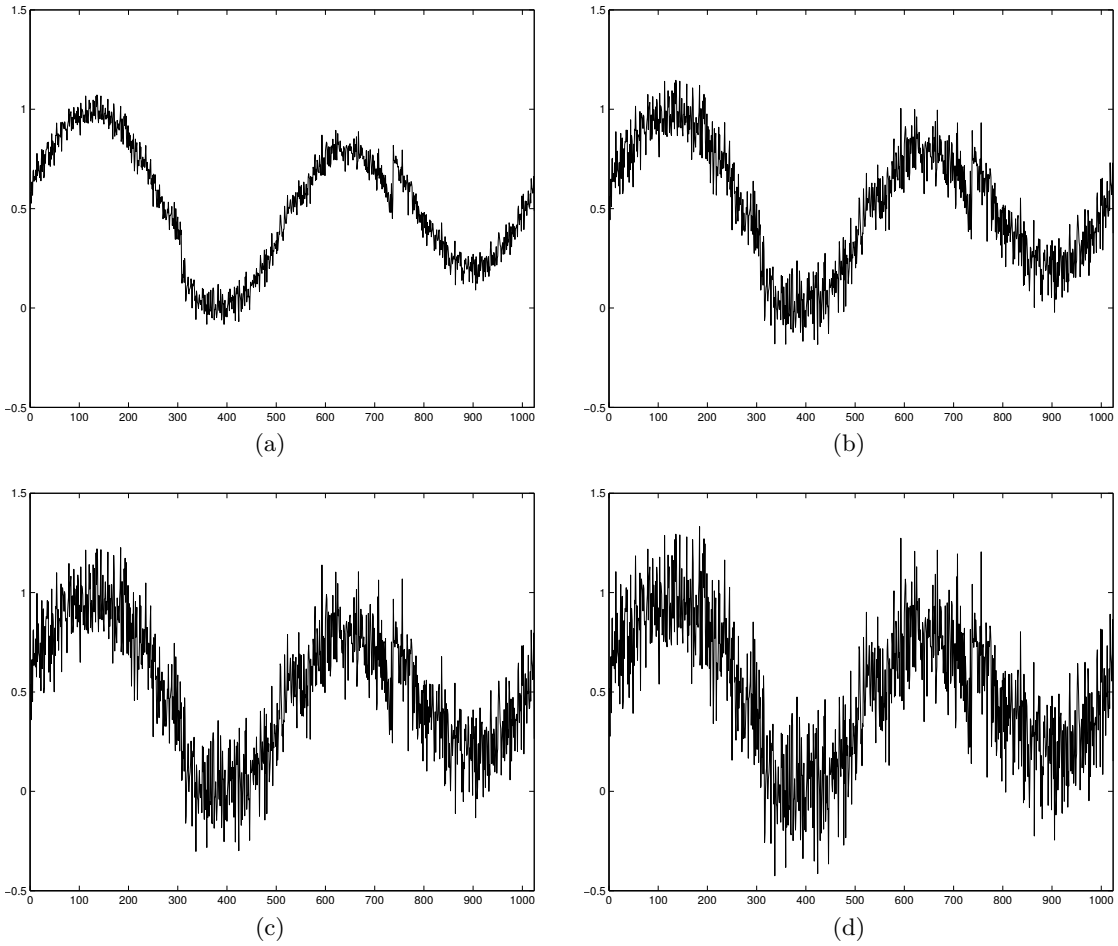


Figure 3.14: *HeaviSine* signal corrupted by noise with σ_ϵ equal to (a) 5%, (b) 10%, (c) 15% and (d) 20% of total signal magnitude

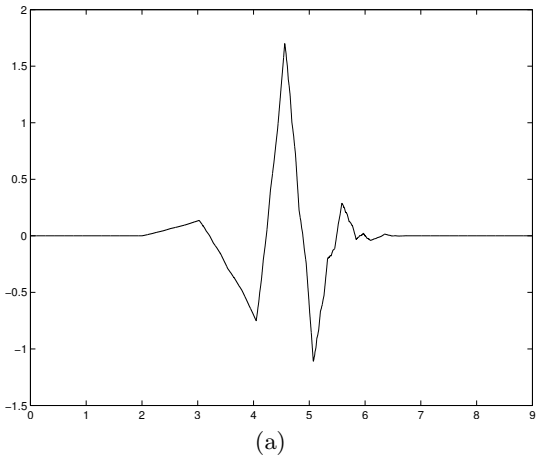


Figure 3.15: Daubechies *Db3* wavelet

support and wavelet function is a piece-wise constant function, so we have a perfect fit of the wavelet shape to the signal.

At moderate noise level, $\sigma\epsilon = 5\%$ of total signal magnitude, ICI-EPL achieves $\approx 28\%$ lower RMSE than the *Haar* wavelet. Noisy and denoised signals are shown in Fig. 3.16. It can be seen that the signal denoised by *Haar* wavelet exert occasional spikes and roughness at edges between the constant values. Constant signal regions are, mostly, efficiently denoised, although some roughness still exists, especially in shorter duration regions. Signal denoised by the ICI-EPL shows almost perfect edge reconstruction and perfect reconstruction of shorter duration constant regions. Performance on longer duration constant regions is similar to the *Haar* wavelet case.

As expected, the best performance is achieved for small Γ parameter values. Larger values result in signal oversmoothing and lower overall quality.

At extreme noise levels, Fig. 3.17, ICI-EPL achieves only $\approx 4\%$ lower RMSE than the *Haar* wavelet. General shape of denoised signals is very similar, but even in this extreme case, ICI-EPL shows much better performance about edges. The edges are still quite sharp, clearly defined, and with no spikes. The optimal Γ parameter value is lower and is just above 0.5. Lowering of the optimal Γ value, as the noise level is increased, will be exhibited in each of the test cases.

Bumps signal

Bumps signal is characterized by a series of bumps (or spires). Except for the bump peaks, the signal can be considered to be smooth. Of the tested conventional wavelets, the *Bior2.2* wavelet performed the best for this type of signal, although *Bior4.4* and *Db3* were also close.

Denoising performance of the ICI-EPL is comparable to the *Bior2.2*. At moderate noise level, the ICI-EPL denoised signal has $\approx 1.5\%$ higher RMSE, with similar visual quality. Fig. 3.18 shows denoising results. If we compare denoised signals with the original *Bumps* signal, showed in Fig. 3.13, we can see that the ICI-EPL slightly better reconstructs bumps than the *Bior2.2* wavelet. It can be most easily spotted when comparing the second bump in the third series. On the other hand, zero level regions in a signal are better reconstructed by the *Bior2.2* wavelet, while the ICI-EPL denoised signal shows some fluctuations and minor spikes in the region.

For this class of signals, it is not possible to achieve desired smoothness in the denoised signal when keeping the Γ parameter at low values. As can be seen in Fig. 3.18 (b), on average, the best results are achieved for $\Gamma > 3$.

The extreme noise case is more difficult to evaluate. In terms of the RMSE, the ICI-EPL scores $\approx 3.8\%$ worse result than the *Bior2.2*, but in terms of the visual quality of denoised signal, a different conclusion can be made.

ICI-EPL performance at zero value regions is indeed lower than the *Bior2.2* performance. There are more oscillations and several unwanted spikes, particularly the ones

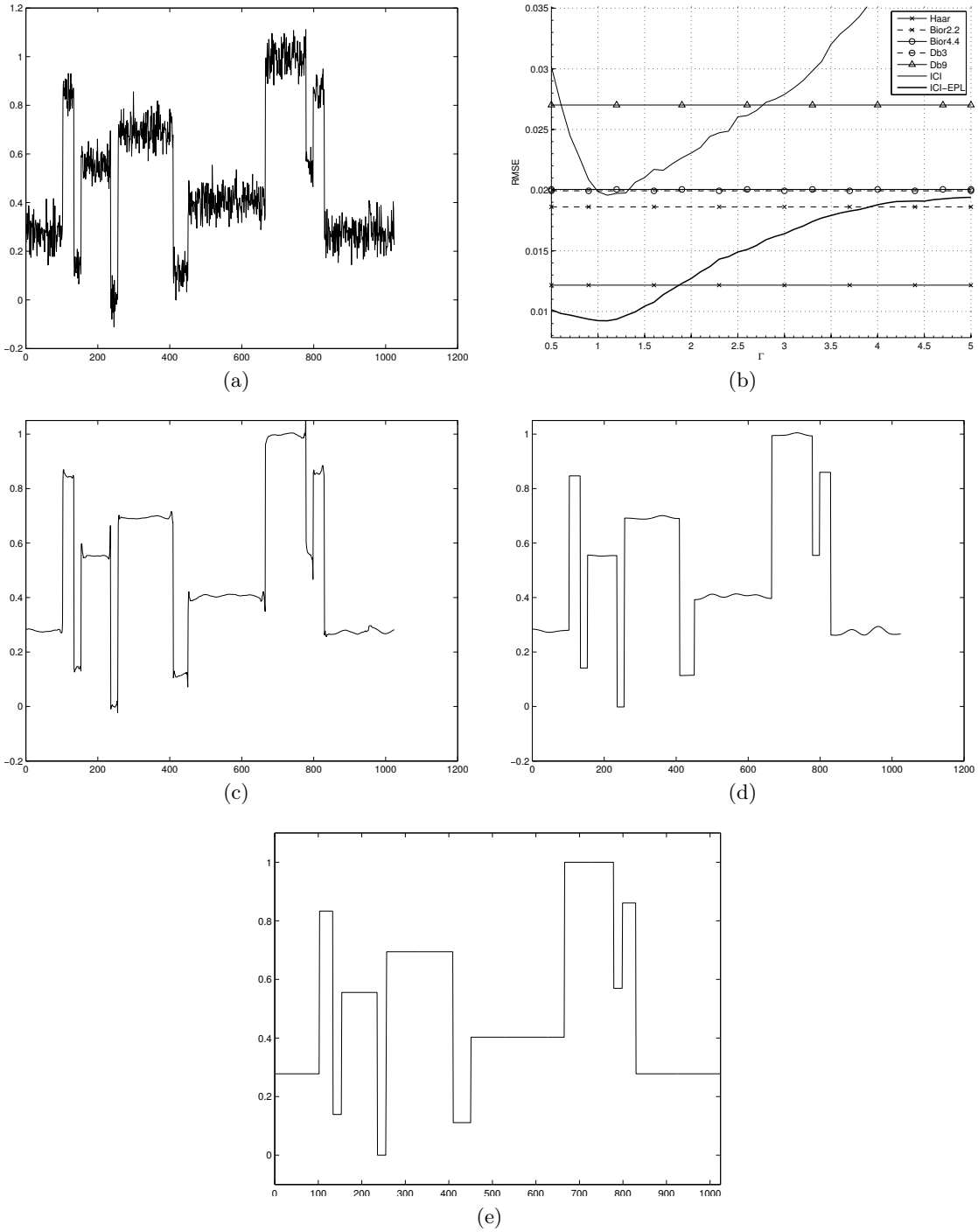


Figure 3.16: *Blocks* signal corrupted by noise with $\sigma_\epsilon = 5\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Haar*) (c), signal denoised by the ICI-EPL (d) and original signal (e)

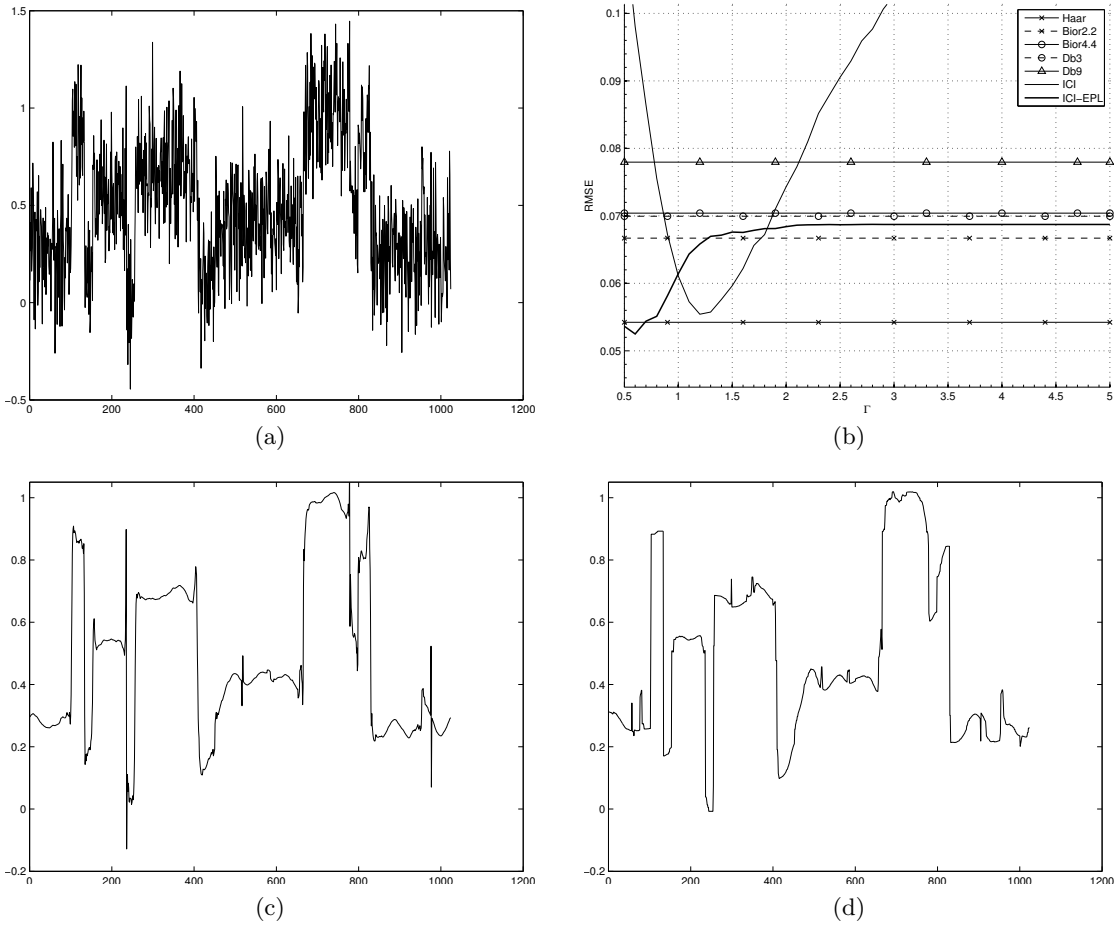


Figure 3.17: *Blocks* signal corrupted by noise with $\sigma_\epsilon = 20\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Haar*) (c) and signal denoised by the ICI-EPL (d)

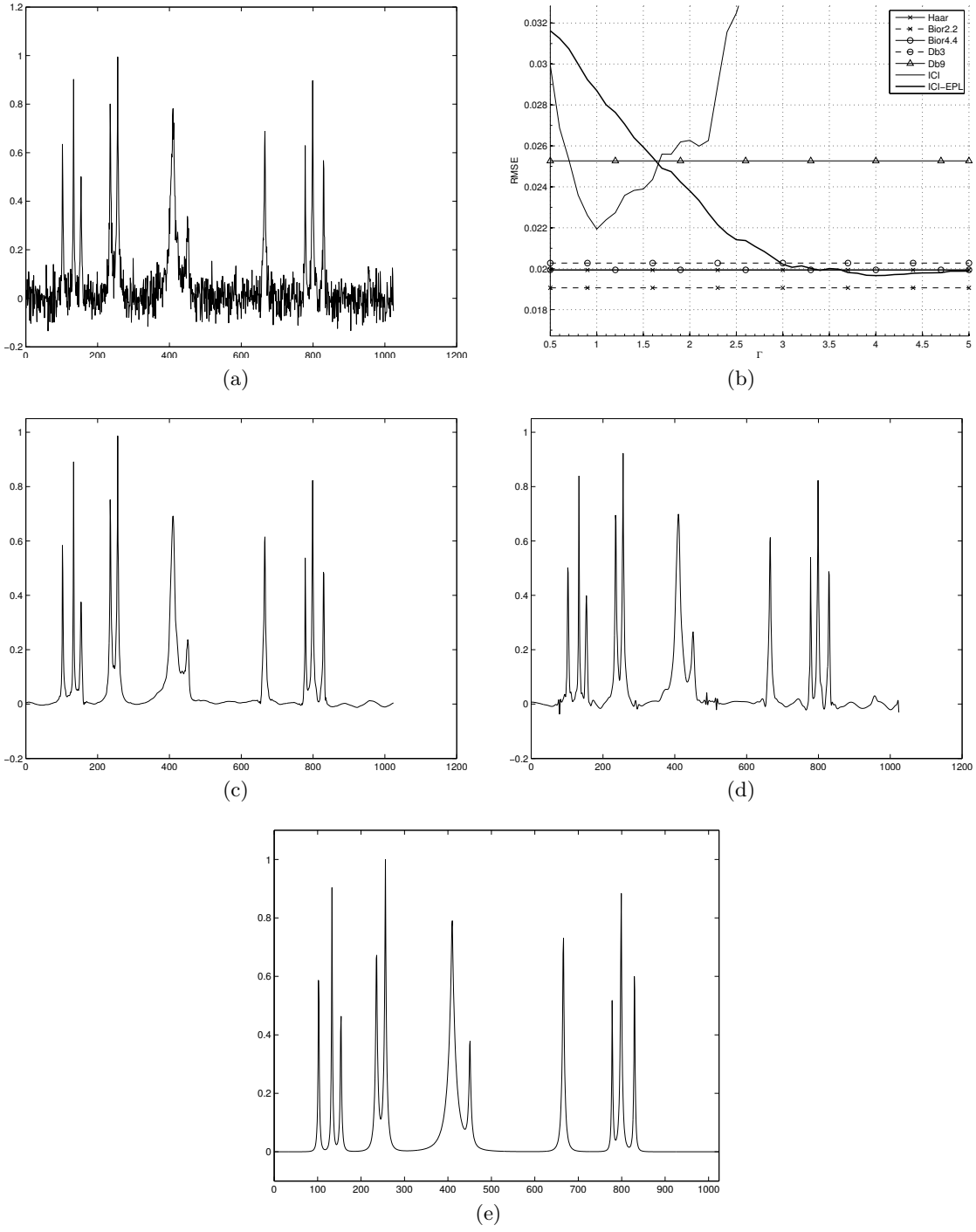


Figure 3.18: *Bumps* signal corrupted by noise with $\sigma_\epsilon = 5\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Bior2.2*) (c), signal denoised by the ICI-EPL (d) and original signal (e)

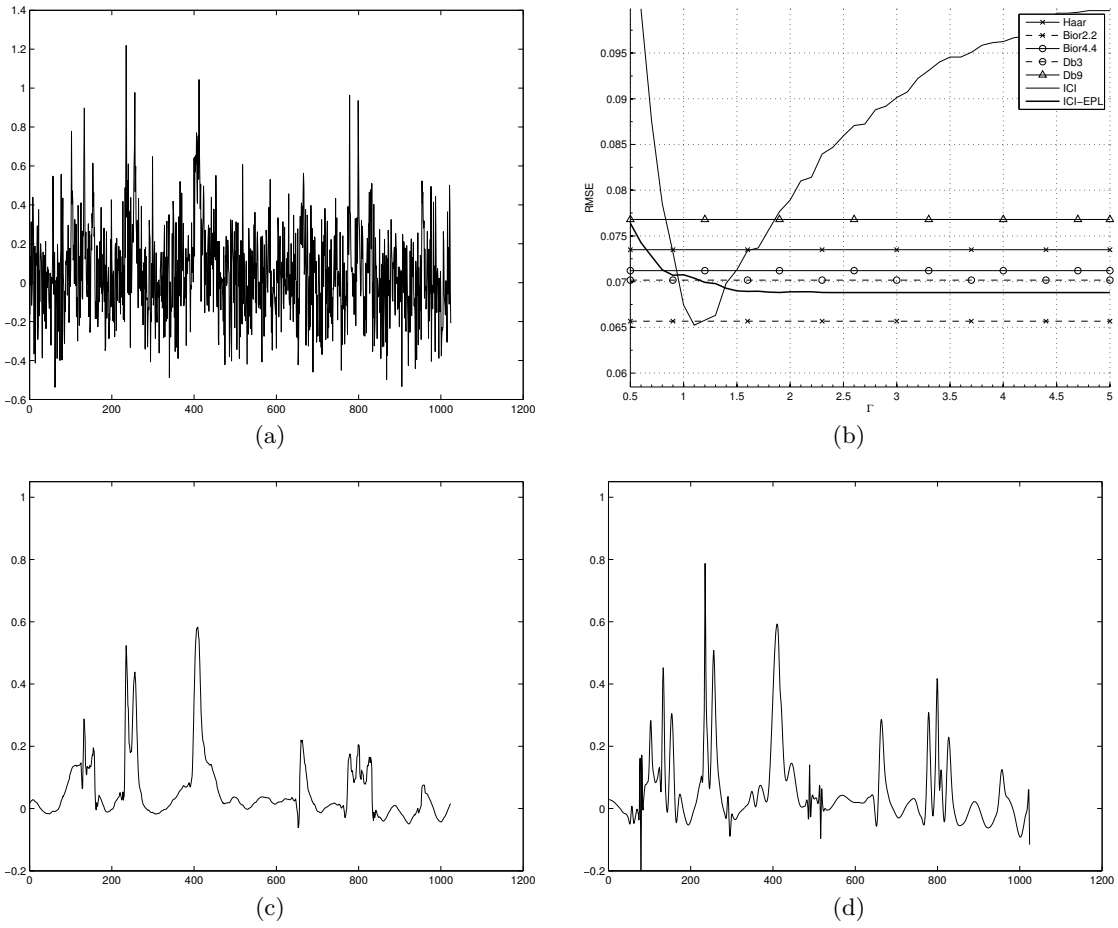


Figure 3.19: *Bumps* signal corrupted by noise with $\sigma_\epsilon = 20\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Bior2.2*) (c) and signal denoised by the ICI-EPL (d)

before the first series of bumps. Still, as can be seen in Fig. 3.19 (d), the false spikes are easily distinguished from the real bumps in a signal. When we compare the bump reconstruction, ICI-EPL performance is far superior. The first series of bumps is badly corrupted in the *Bior2.2* case, while the ICI-EPL successfully conserved all three bumps in the series. Similar observation can be made for the fourth series of bumps. Although the *Bior2.2* performed better than for the first series, it is still inferior to the ICI-EPL. Also, bump peaks in the ICI-EPL denoised image are much higher, and closer to the actual peaks in the original signal.

Following above considerations and the fact that the bumps in the *Bumps* signal, and not zero value regions, carry most of the useful signal information, we conclude that the ICI-EPL performed significantly better in the extreme noise case, although the RMSE does not suggest the same.

Doppler signal

Amplitude and frequency changes of the *Doppler* signal are challenging for denoising algorithms. Since lowest signal amplitudes are associated with highest signal frequencies, it is inevitable that the noise corruption will cause masking of actual signal features in such regions. As the frequency declines and amplitude rises, the signal becomes very smooth. The optimal wavelet to be used for denoising the *Doppler* signal, with respect to RMSE, is the *Bior4.4* wavelet.

At moderate noise level, performance of the ICI-EPL and the *Bior4.4* are comparable. ICI-EPL achieves better performance for larger Γ parameter values, with the best result being at the largest value, $\Gamma = 5$. RMSE is $\approx 1.8\%$ higher than for the *Bior4.4*, with visual quality of denoised signals (Fig. 3.20) also being comparable. The *Bior4.4* has small performance edge in the highest frequency region, while the ICI-EPL provides smoother signal in the lower frequency region.

Results are similar in the extreme noise case. RMSE of the ICI-EPL is now $\approx 1.4\%$ lower, and visual quality (Fig. 3.21) is again comparable. Because of the high noise level, wider high frequency region is flattened in both denoised signals. Again, ICI-EPL denoised signal is smoother, throughout the signal duration, except near the end, where the *Bior4.4* performs slightly better.

For both noise levels results of the optimal conventional wavelet transform and the ICI-EPL are comparable in terms of the both, RMSE and the visual quality.

HeaviSine signal

This is another signal with conflicting requirements. Except for the two discontinuities, the signal is completely smooth. As such, the *Bior2.2* wavelet resulted in the smallest RMSE for conventional wavelets.

At moderate noise level, the ICI-EPL performs better than the *Bior2.2*. It yielded denoised image with $\approx 4.5\%$ lower RMSE and much better visual quality. As can

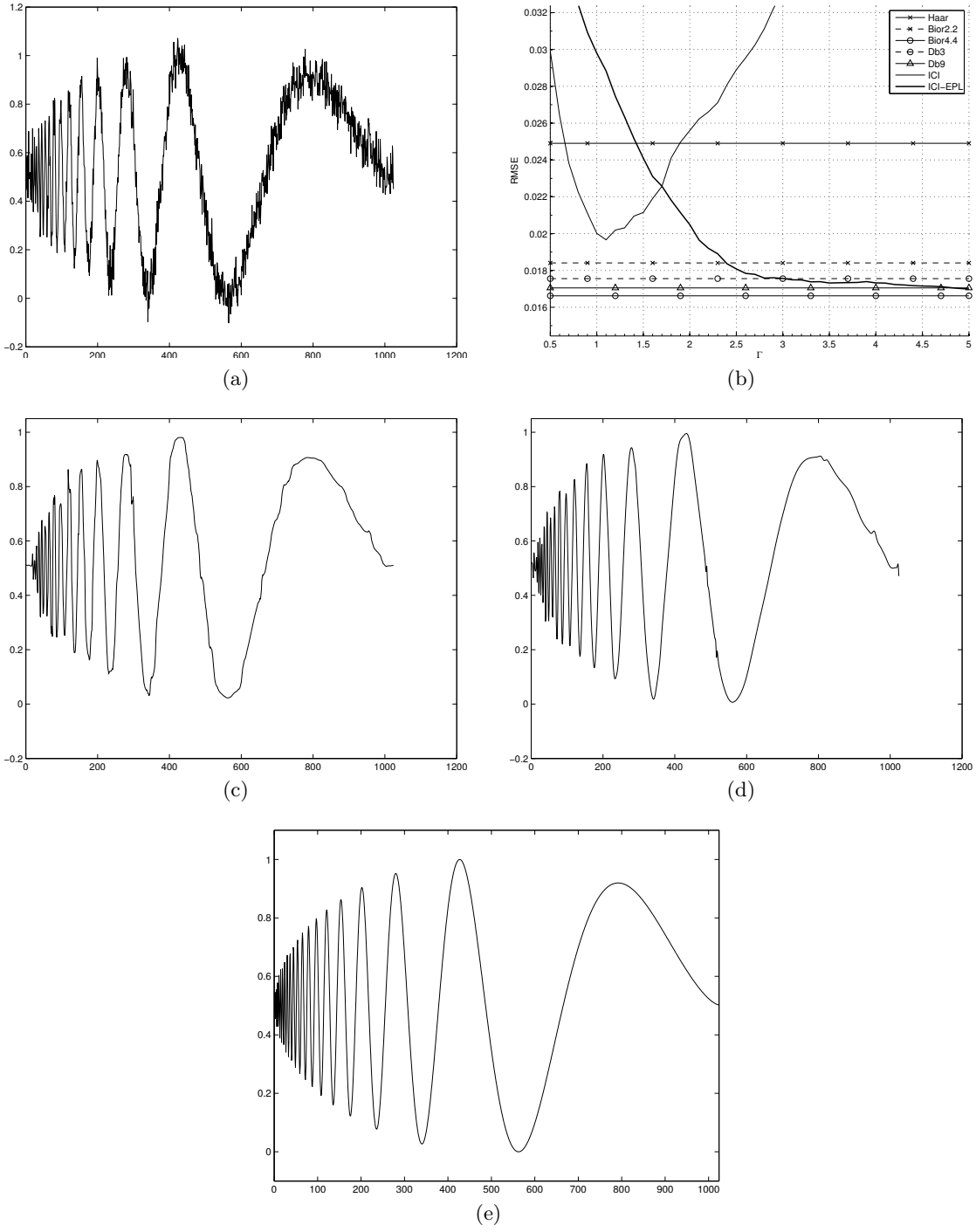


Figure 3.20: *Doppler* signal corrupted by noise with $\sigma_\epsilon = 5\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Bior4.4*) (c), signal denoised by the ICI-EPL (d) and original signal (e)

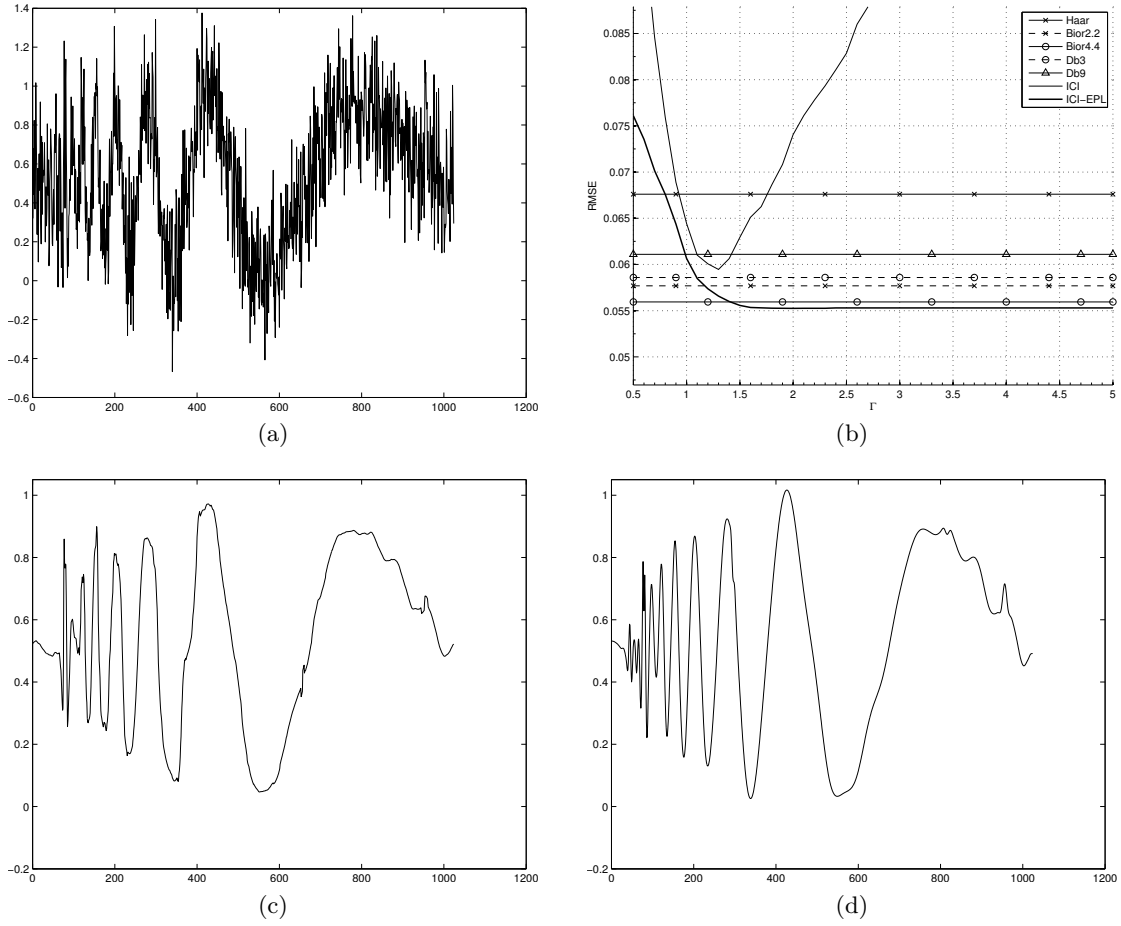


Figure 3.21: *Doppler* signal corrupted by noise with $\sigma_\epsilon = 20\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Bior4.4*) (c) and signal denoised by the ICI-EPL (d)

be seen in Fig. 3.22, the bulk of the signal is very smooth, while the discontinuities are accurately reconstructed. The *Bior2.2* denoised signal also features reasonable discontinuity reconstruction, but is not as smooth as the ICI-EPL denoised signal.

On average, the best results for ICI-EPL are achieved for smaller Γ parameter values, with minimum being at $\Gamma = 1.3$. Nature of the signal allows for such small values, as the adaptive algorithm can easily recognize signal discontinuities, while still allowing for longer filters to be used for smooth regions.

At extreme noise level, the *Bior2.2* is no longer the optimal choice, and the *Haar* wavelet performs better. In terms of the RMSE, the *Haar* is superior to the ICI-EPL and results in $\approx 11\%$ lower error. The difference is not confirmed in the visual quality inspection. As shown in Fig. 3.23, the first discontinuity is almost completely lost in both denoised signals, while the second one is equally attenuated. The biggest difference between the signals is in the troughs and near the signal end, where the *Haar* wavelet performed better.

We can conclude that the transforms yield comparable results, with small advantage going to the *Haar* wavelet.

Piece-Polynomial signal

Piece-Polynomial is a piece-wise polynomial signal, with sharp edges (discontinuities) between the pieces. The only tested conventional wavelet transform that produced reasonable result is the *Haar* wavelet. All the other transforms perform inadequately about edges and show significantly worse RMSE and visual quality.

At moderate noise level, the signal can be denoised very efficiently using the ICI-EPL. Signal denoised by the *Haar* wavelet yields $\approx 21.5\%$ higher RMSE than in the ICI-EPL case. Huge difference in RMSE can, also, be confirmed by visual quality evaluation (Fig. 3.24). Larger polynomial pieces are comparable between the two denoised signal, but difference in shorter pieces and about edges is distinctive. Edges in the ICI-EPL denoised signal are completely preserved and accurately reconstructed. It is not true for the *Haar* case, in which many spikes are present about edges. The *Haar* wavelet performance about edges also has influence on reconstruction of smaller polynomial pieces, which is much better in the ICI-EPL case.

Like for the *Blocks* signal, lower parameter Γ values are preferred for denoising the *Piece-Polynomial* signal. On average, the best RMSE is achieved about $\Gamma = 1$.

In extreme noise case the ICI-EPL do not show the same advantage. In terms of the RMSE, it shows only $\approx 1\%$ better performance. When comparing visual quality of denoised signals (Fig. 3.25), we can see that the signals are very similar. The only major difference is the existence of several spikes in the *Haar* case, which is the main cause for the difference in RMSE. ICI-EPL does seem to have a very slight advantage over the *Haar* wavelet but, in general, the two transforms produce comparable results.

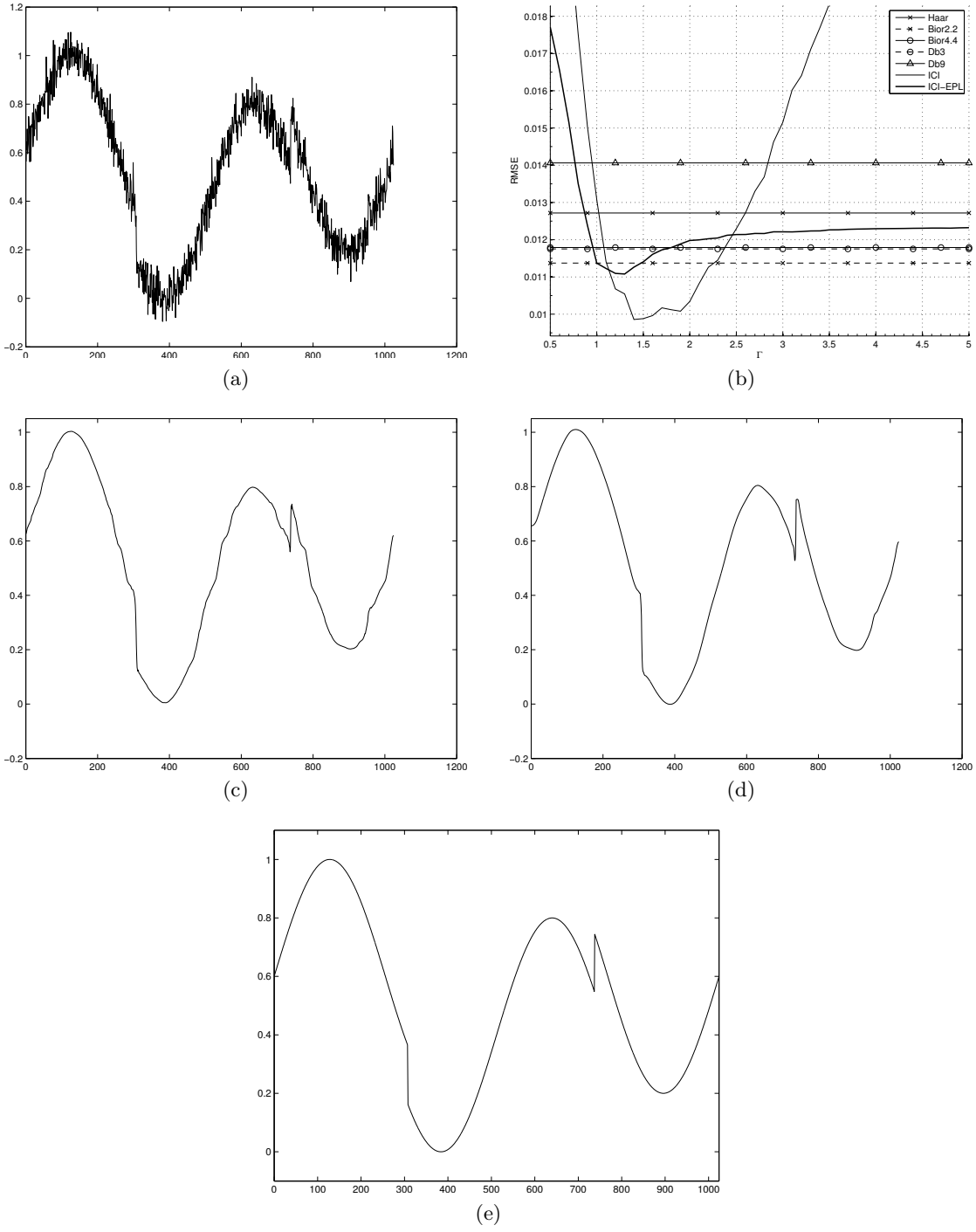


Figure 3.22: *HeaviSine* signal corrupted by noise with $\sigma_\epsilon = 5\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Bior2.2*) (c), signal denoised by the ICI-EPL (d) and original signal (e)

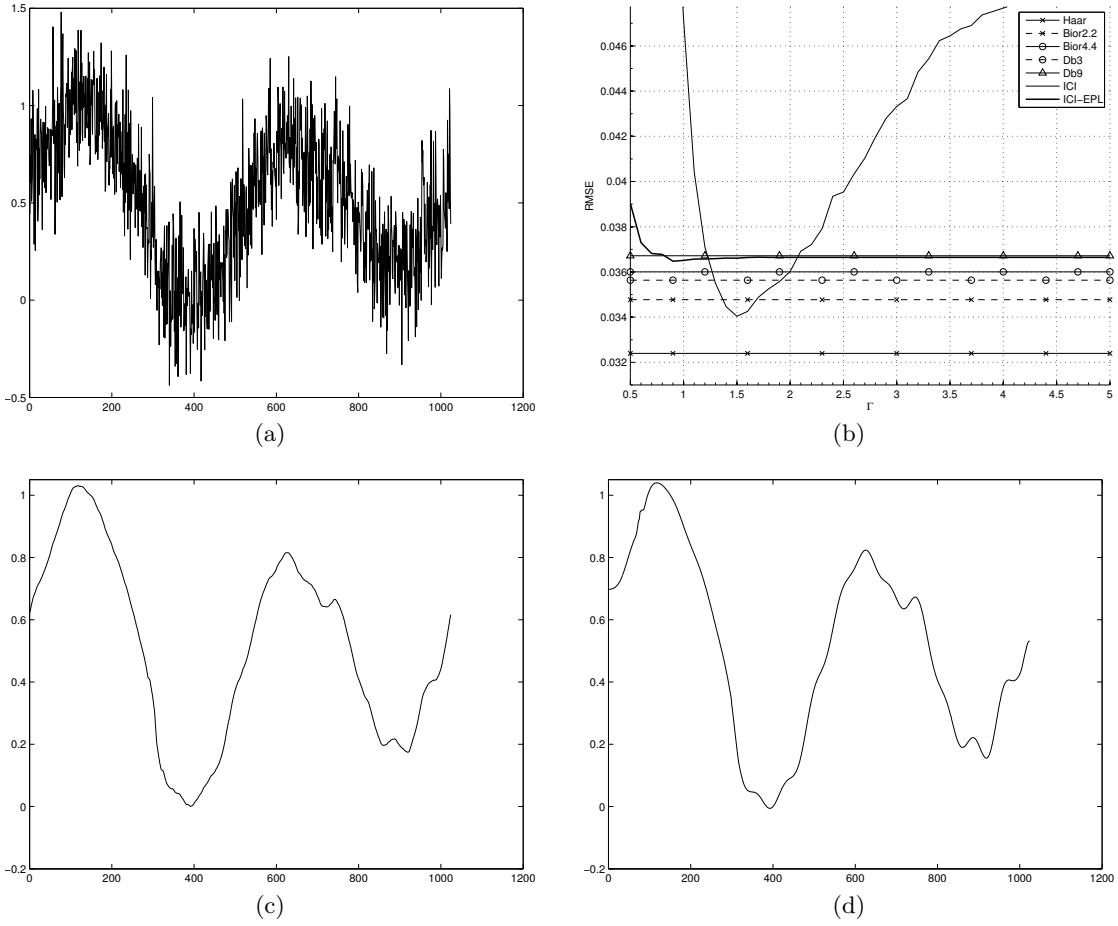


Figure 3.23: *HeaviSine* signal corrupted by noise with $\sigma_\epsilon = 20\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Haar*) (c) and signal denoised by the ICI-EPL (d)

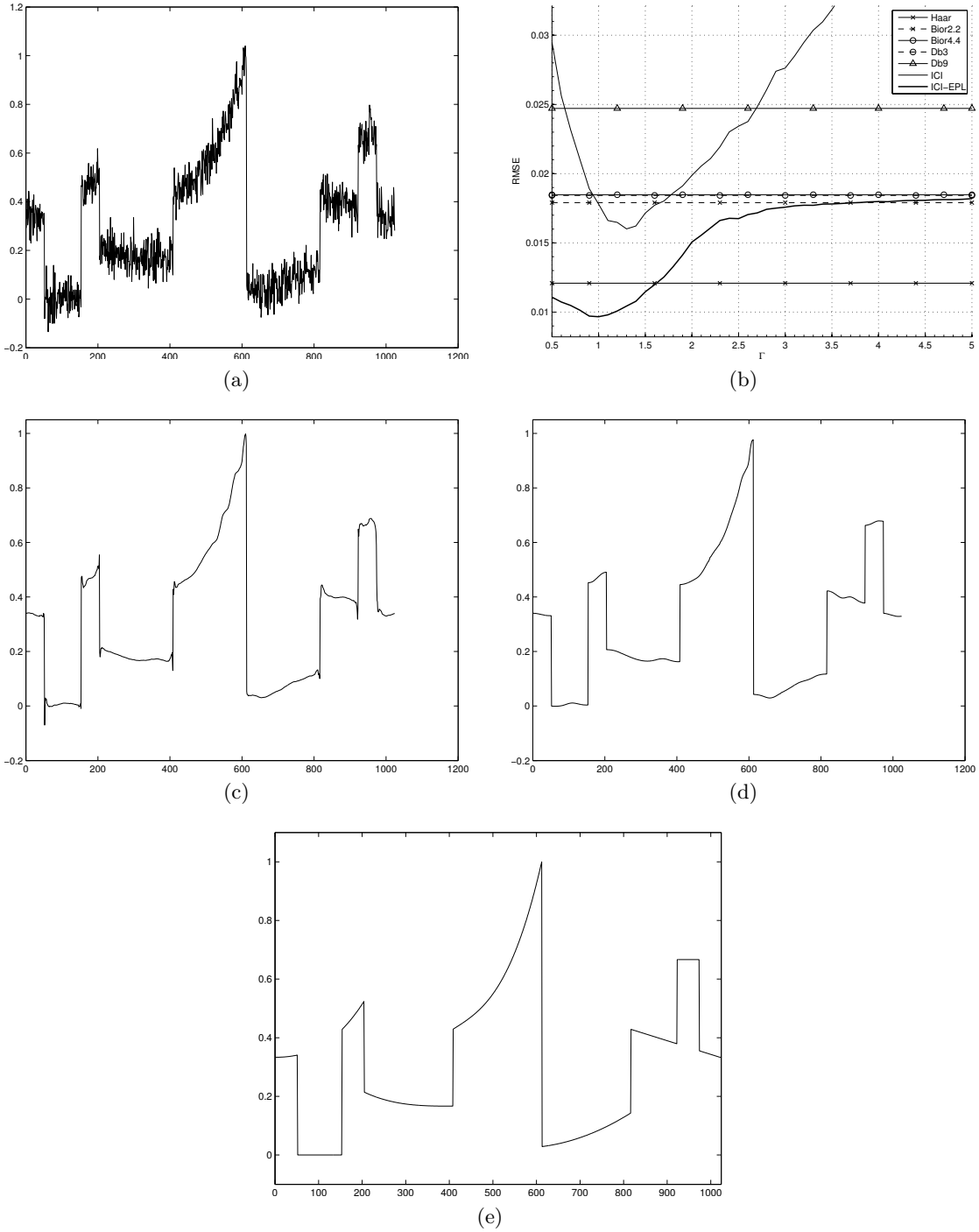


Figure 3.24: *Piece-Polynomial* signal corrupted by noise with $\sigma_\epsilon = 5\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Haar*) (c), signal denoised by the ICI-EPL (d) and original signal (e)

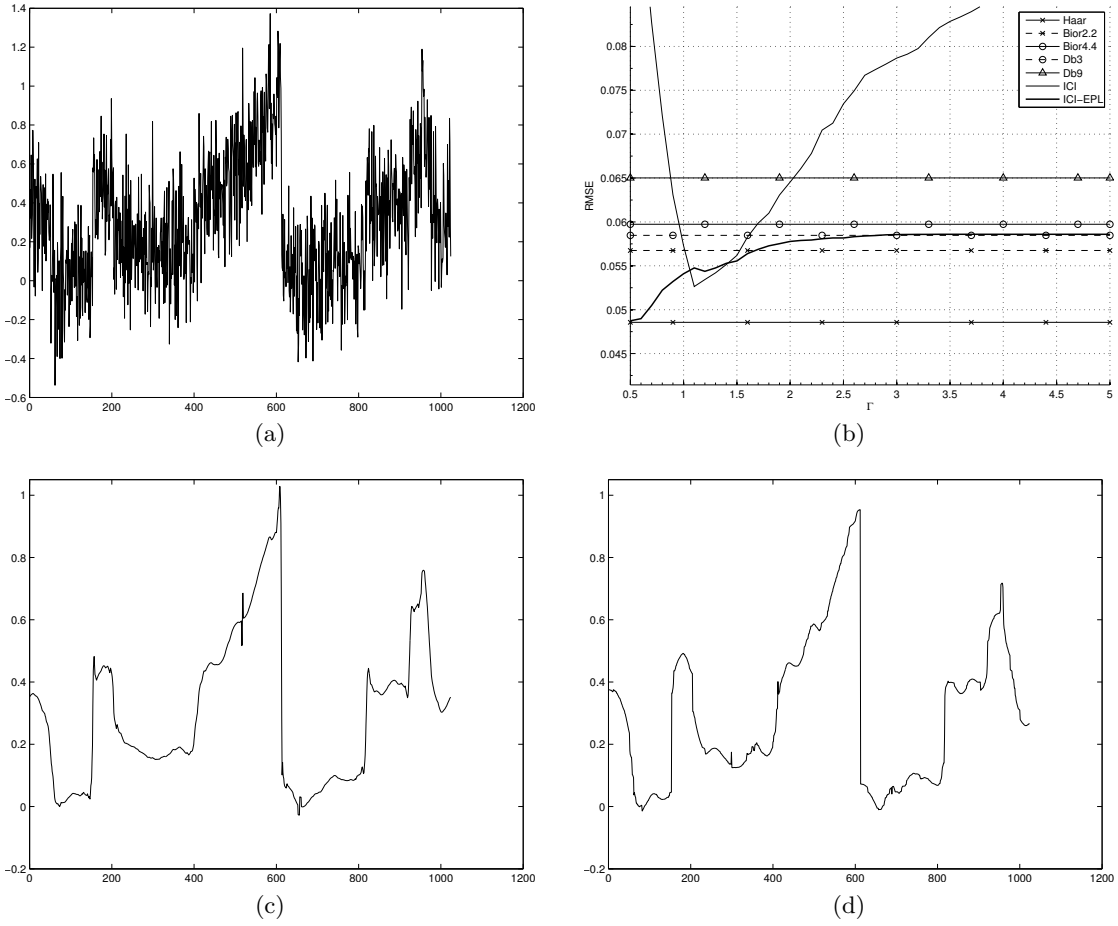


Figure 3.25: *Piece-Polynomial* signal corrupted by noise with $\sigma_\epsilon = 20\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Haar*) (c) and signal denoised by the ICI-EPL (d)

***Piece-Regular* signal**

The last signal to be analyzed is the *Piece-Regular* signal. It is characterized by many singularities, which take form of either bumps or discontinuities with sharp edges. *Db3* is the wavelet which scored the lowest RMSE of the conventional wavelets tested.

As we already showed, in case signal contains sharp edges, the ICI-EPL denoising scheme always shows definite advantage over conventional wavelet transforms . The same can be noted for the *Piece-Regular* signal. At moderate noise level, the ICI-EPL yields $\approx 18\%$ lower RMSE than the *Db3*. If we look at the visual quality of denoised signals (Fig. 3.26), we can see that the overall quality of the signals is mostly comparable, except about edges. The *Db3* wavelet denoised signal features spikes at many edges, while the ICI-EPL succeeded to efficiently recover signal information and accurately reconstruct edges between regular pieces in the signal.

At extreme noise level, difference between the two denoising methods is again diminished. For the given noise level, the *Haar* wavelet performs better than the *Db3* and results in $\approx 1\%$ lower RMSE than the ICI-EPL. As suggested by the RMSE, both of the transforms produce comparable results, even in the terms of visual quality. However, as can be seen in Fig. 3.27, several large spikes exist in the *Haar* denoised signal, which are not present in the ICI-EPL denoised signal. Therefore, we favor the ICI-EPL over the *Haar* wavelet for denoising the *Piece-Regular* signal, even at the extreme noise levels.

3.6 Conclusion

The ICI-EPL adaptive lifting scheme for signal denoising was proposed. The goal was to develop an edge preserving transform, such that the good properties of conventional wavelet transforms are retained, but performance about edges in a signal is greatly improved. The adaptation is carried out by selecting appropriate wavelet basis, on each scale and for each signal sample, independently.

The intersection of confidence intervals (ICI) rule is the core of the adaptive algorithm. It is used in the predict step of the lifting scheme to determine the neighborhood of samples, which share the same local properties as the sample whose value is being predicted. This way, longer lifting filters are used in the smoother signal regions, while shorter wavelets are used in the higher frequency regions. As a result, the transform efficiently reconstructs both, the edges in a signal and the smooth signal regions.

Extensive experiments were performed to evaluate efficiency of the ICI-EPL. Different classes of signals were used, sampled at several frequencies and with various levels of additive noise. We showed that in many cases, the ICI-EPL easily outperforms the ICI smoothing method and all the tested conventional wavelet transforms. When this is not the case, it produces results which are comparable to the transform yielding the best results for a given signal. The edge preserving nature can be clearly seen in denoised signals, as edges are efficiently reconstructed, without oversmoothing, or exhibiting pseudo-Gibbs

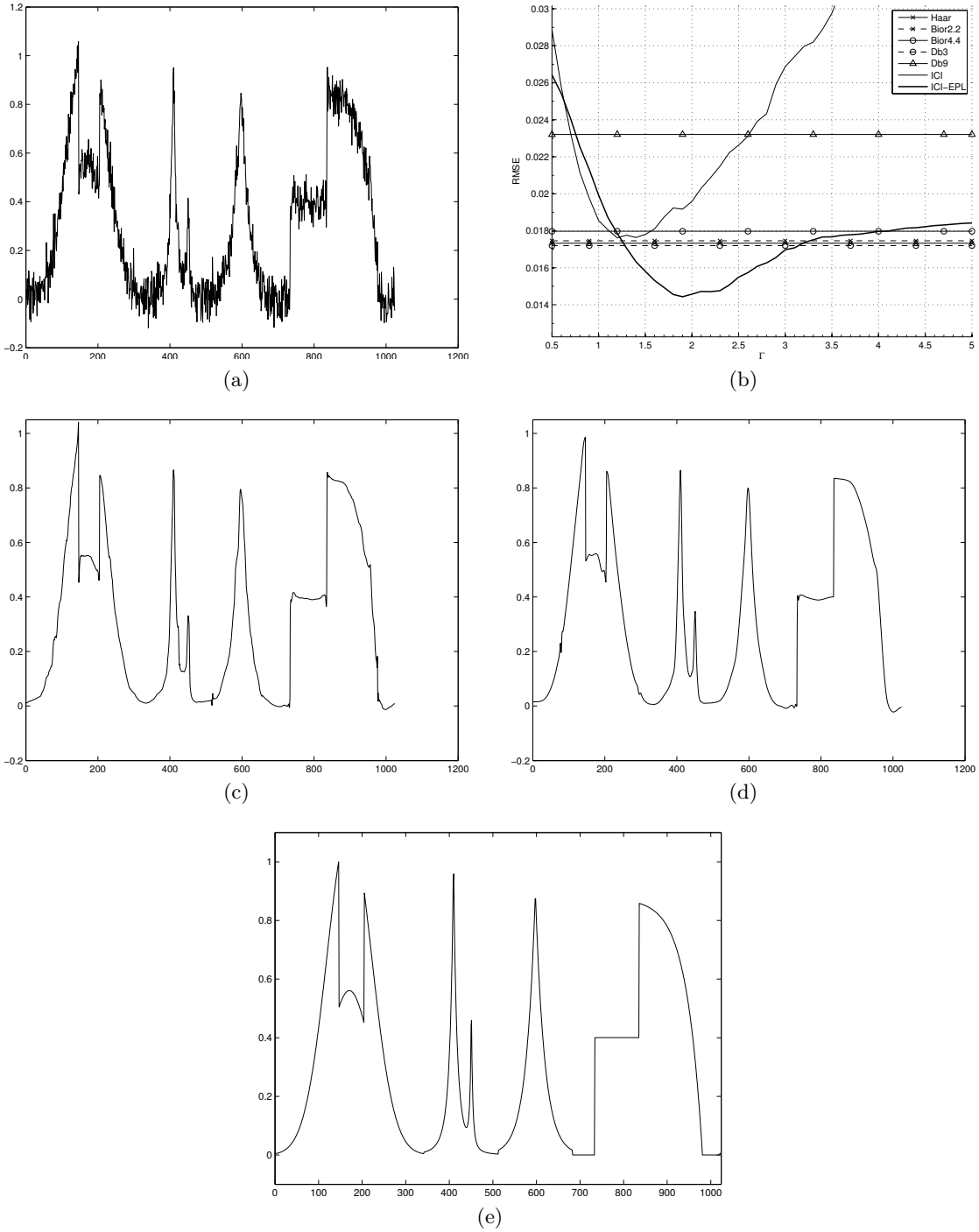


Figure 3.26: *Piece-Regular* signal corrupted by noise with $\sigma_\epsilon = 5\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Db3*) (c), signal denoised by the ICI-EPL (d) and original signal (e)

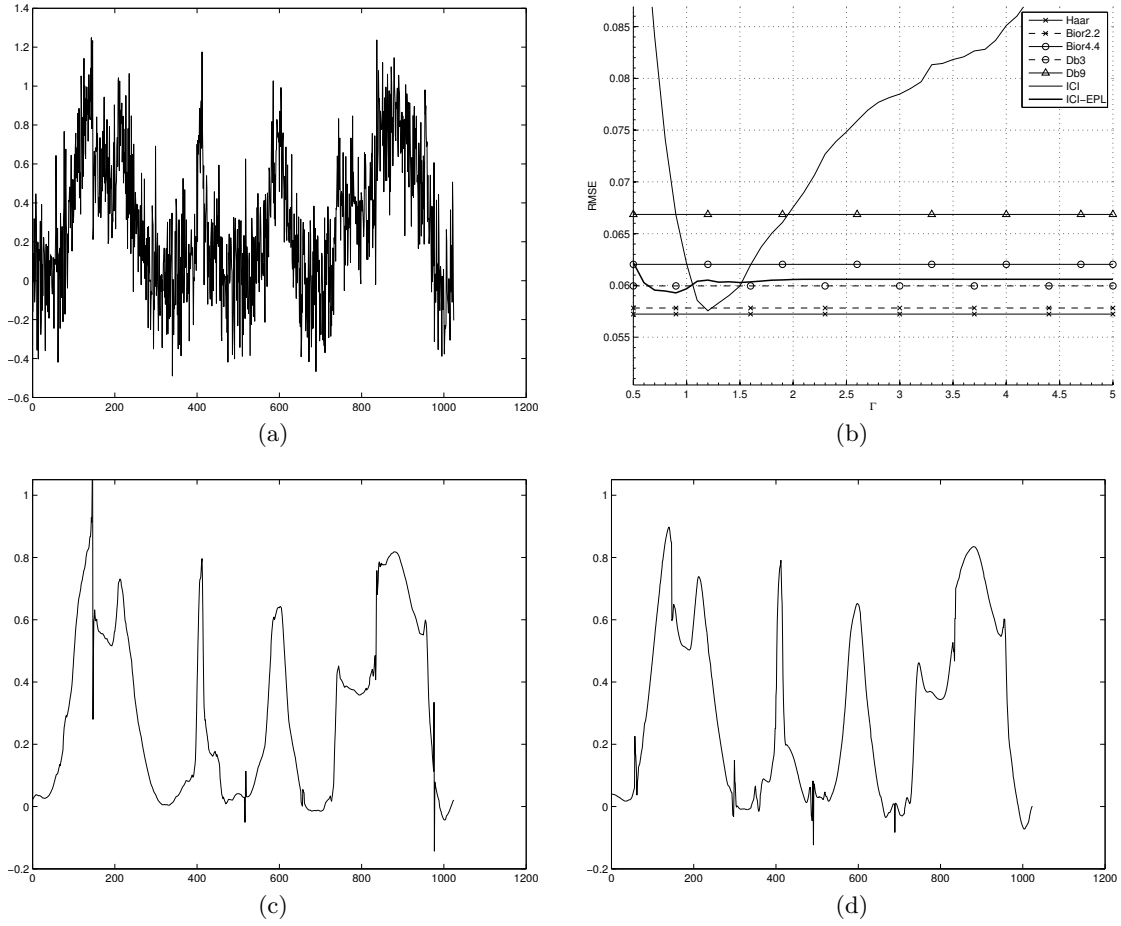


Figure 3.27: *Piece-Regular* signal corrupted by noise with $\sigma_\epsilon = 20\%$ of total signal magnitude (a), average RMSE for various denoising methods (b) signal denoised by optimal wavelet transform (*Haar*) (c) and signal denoised by the ICI-EPL (d)

oscillations.

The main deficiency of the ICI-EPL is its dependence on the proper choice of the ICI Γ parameter value, as false selection can have detrimental influence on denoising performance. We will try to solve it in the next chapter, by proposing statistical method for automated parameter Γ selection.

Chapter 4

Determining the Γ Parameter Value

It was shown in Chapter 3, that the ICI-EPL performance in signal denoising can vary significantly, depending on the chosen value of the parameter Γ. Each signal assumes its own optimal Γ value and, if correctly selected, the transform will perform very well. If the optimal value is missed, the transform performance may downgrade significantly. Unfortunately, there is no direct analytical way of finding the value. Katkovnik in its original work (Katkovnik, 1999) and later in (Katkovnik, Egiazarian, and Astola, 2002) proposed cross-validation as a tool for determining the optimal value. A predefined set of N_G values, was used to perform denoising using the ICI denoising scheme. After the denoising, a cross-validation loss function is constructed using the N_G denoised signals and associated kernel estimators. Chosen value for the Γ is a value for which the cross-validation loss function has a minimum. To achieve satisfying results, it is necessary for the kernel estimators to have large maximum support lengths. The condition is fulfilled in original Katkovnik work, but in the ICI-EPL case, filters with compact support are used, so a different approach is necessary.

In this chapter, we propose two methods for the parameter Γ value selection, which are both based on the statistical distribution of wavelet coefficients. We show that the distribution is linked to the efficiency of the adaptive algorithm, and we use that connection to decide the Γ value for a given signal.

4.1 Distribution of Wavelet Coefficients

At each decomposition level of a wavelet transform, lowpass and highpass subbands are generated. As presented in greater detail in Chapter 2, the lowpass subband contains signal average, i.e., the coarse approximation of the signal, while the highpass subband contains signal details. We can assume the highpass subband coefficients (wavelet coefficients) to be realizations of independent and identically distributed random variables. Distribution

of the variables carry important information about appropriateness of a wavelet transform for a given signal. Extensive research has been carried out in analyzing and modeling the distribution for image processing applications, for instance (Lam, 2004; Mallat, 1989c; Wang, Li, Xuan, Lo, and Mun, 1997). Their main conclusions are also valid for the 1-D signals. Generally, authors agree that the distribution of wavelet coefficients can be modeled as a generalized Gaussian distribution (GGD). However, it is not possible to use the same GGD to model distribution of any set of wavelet coefficients, because the coefficients distribution is neither unique nor constant. It changes at each scale (decomposition level) and it is influenced by the type of signal being decomposed, the type and level of additive noise, and the chosen wavelet basis. In order to devise the parameter Γ selection methods, we need to consider each of the factors influencing the highpass subband coefficients distribution.

For the beginning, let us recall that we assumed the given signal $z(x)$ to be a noisy observation of the original signal $y(x)$:

$$z(x) = y(x) + \epsilon(x). \tag{4.1}$$

Since wavelet transform is a linear transform, it is true that the actual transform of the $z(x)$ consists of two separate components:

$$\mathcal{W}(z(x)) = \mathcal{W}(y(x)) + \mathcal{W}(\epsilon(x)). \tag{4.2}$$

The same is also true for distributions of wavelet coefficients. They are influenced by both, the coefficients originating from the original signal and the coefficients originating from noise. As wavelet decomposition progresses, the influence of the two sources on wavelet coefficients changes. In effect, on different decomposition levels, distribution of wavelet coefficients will be more impacted by either the noise component or the original signal component of $z(x)$. Because of (4.2), we can study each of the components separately.

4.1.1 Influence of Scale

The effect of scale on distribution of wavelet coefficients is depicted in Fig. 4.1. It shows the distributions at 5 decomposition levels. Signals were decomposed using the *Bior2.2* as a wavelet basis. It can be seen that in most cases the distribution at level 1 closely resembles the Gaussian probability density function (PDF). Further down the decomposition tree, the PDF steadily narrows and its peak rises. The exception to this is the *Bumps* signal, Fig. 4.1(b), for which the PDF at level 5 has higher peak than at the level 4. The exception can be attributed to the influence of other factors, as will be seen later. The same is true for the *Heavisine* signal, Fig. 4.1(d), for which the difference between PDFs at different decomposition levels is not that definite, as in the other cases.

The dynamics of the wavelet coefficients PDF, as seen in Fig. 4.1, is expected. All of

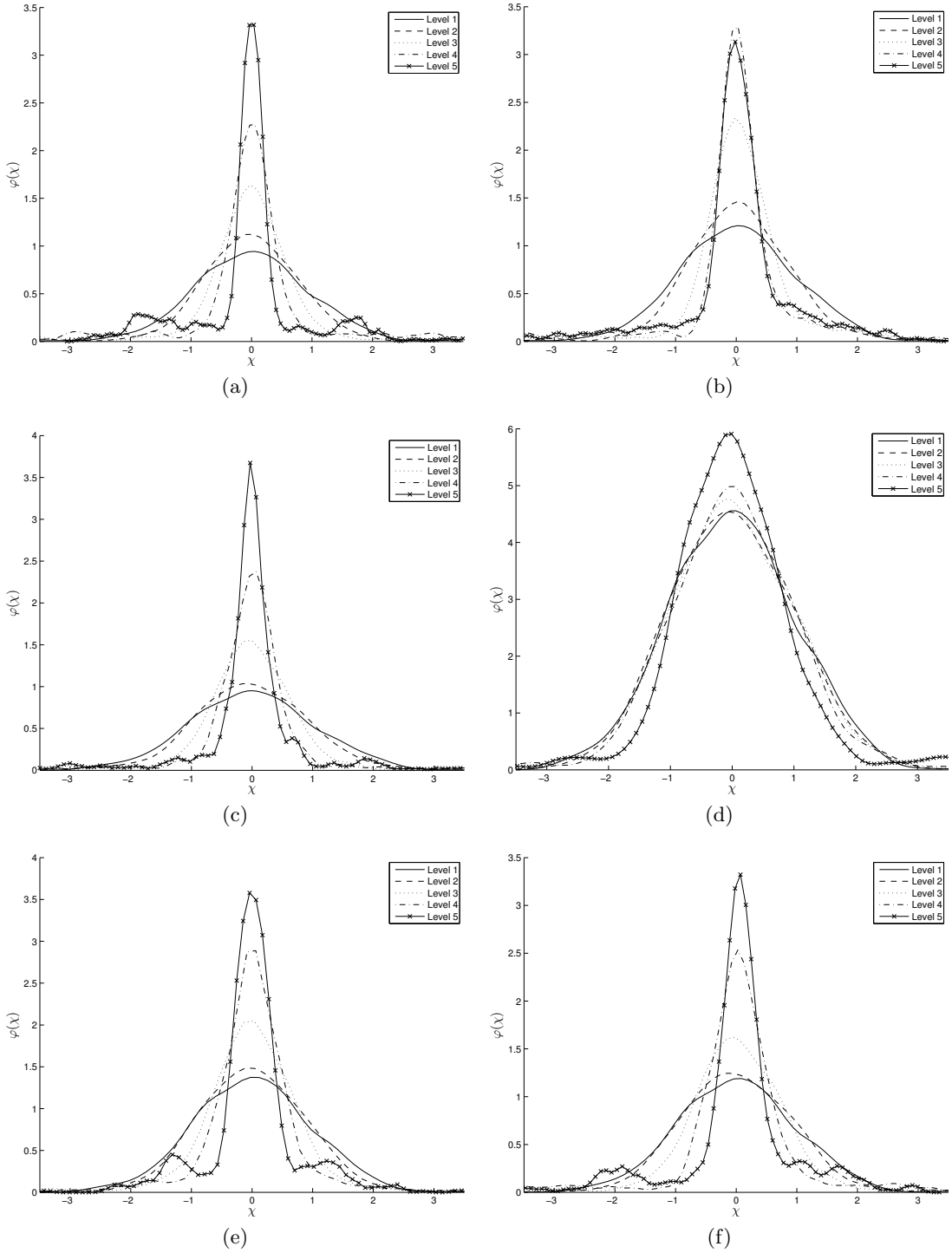


Figure 4.1: Distribution of wavelet coefficients at each decomposition level, for signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f). Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal

the original test signals considered belong to a class of signals with power spectral density of the form:

$$S(f) \propto \frac{1}{f^{\alpha}}, \qquad \alpha > 0, \tag{4.3}$$

which means that signals are mostly constituted of low frequencies. On the other hand, additive white Gaussian noise contains equally all frequencies. As a result, the highpass filter at the first decomposition level filters out most of the original signal, and, normally distributed, wavelet coefficients almost entirely originate from the additive noise. This can be confirmed in Fig. 4.2, which shows PDFs of wavelet coefficients at the first decomposition level. Three different wavelet basis were used, namely, *Haar*, *Bior2.2*, *Bior4.4*. For comparison, the Gaussian PDF is also shown. Obviously, changing the wavelet or the signal do not have considerable influence on the PDF. For all the test signals and the three different wavelet basis the PDF is very close to the Gaussian PDF. Therefore, it is difficult to find a connection between the coefficient distribution at the first decomposition level, and suitability of a given wavelet basis for a specific signal.

As we go down the decomposition tree, the highpass filter passes progressively narrower and lower frequency bands. On each successive scale, impact of noise on wavelet coefficients fades and coefficients mostly originate from the original signal. Because of the energy packing property of wavelet transforms, most of the signal information will be contained in a small number of relatively large wavelet coefficients. In effect, the distribution of coefficients is expected to be narrow and peaked about zero. This is clearly visible in Fig. 4.3, in which the PDFs at the last decomposition level are shown.

In this case, there is a significant difference between the PDFs from the three wavelet basis. More importantly, there is a clear connection between the PDFs shape and the suitability of a given wavelet basis for a specific signal. It leads us to a hypothesis that will be shown valid for all test signals and will be central for the proposed method for the Γ parameter selection.:

> The better the wavelet basis fits local signal properties, the narrower will be the shape of the wavelet coefficients PDF at the last decomposition level, and the higher its peak.

The two most exemplifying cases of the hypothesis validity are the PDFs for the *Blocks* and the *Doppler* signals. The signals itself are shown in Fig. 3.13(a), and Fig. 3.13(c), respectively. *Blocks* signal is a piece-wise constant signal, characterized by sharp edges between the constant pieces. As such, the ideal wavelet basis for the signal is the *Haar* wavelet (Fig. 3.5 (a)). Its shape allows for an excellent reconstruction of sharp edges in the signal. The PDFs of wavelet coefficient at the last decomposition level are shown in Fig. 4.3(a). The PDF for the *Haar* transform has decisively higher peak and is narrower than the other two PDFs. It is followed by the *Bior2.2* and finally the *Bior4.4*.

The PDFs for the *Doppler* signal, Fig. 4.3(c), are subject to a similar interpretation. The signal is characterized by lower frequencies and smooth regions which can be efficiently

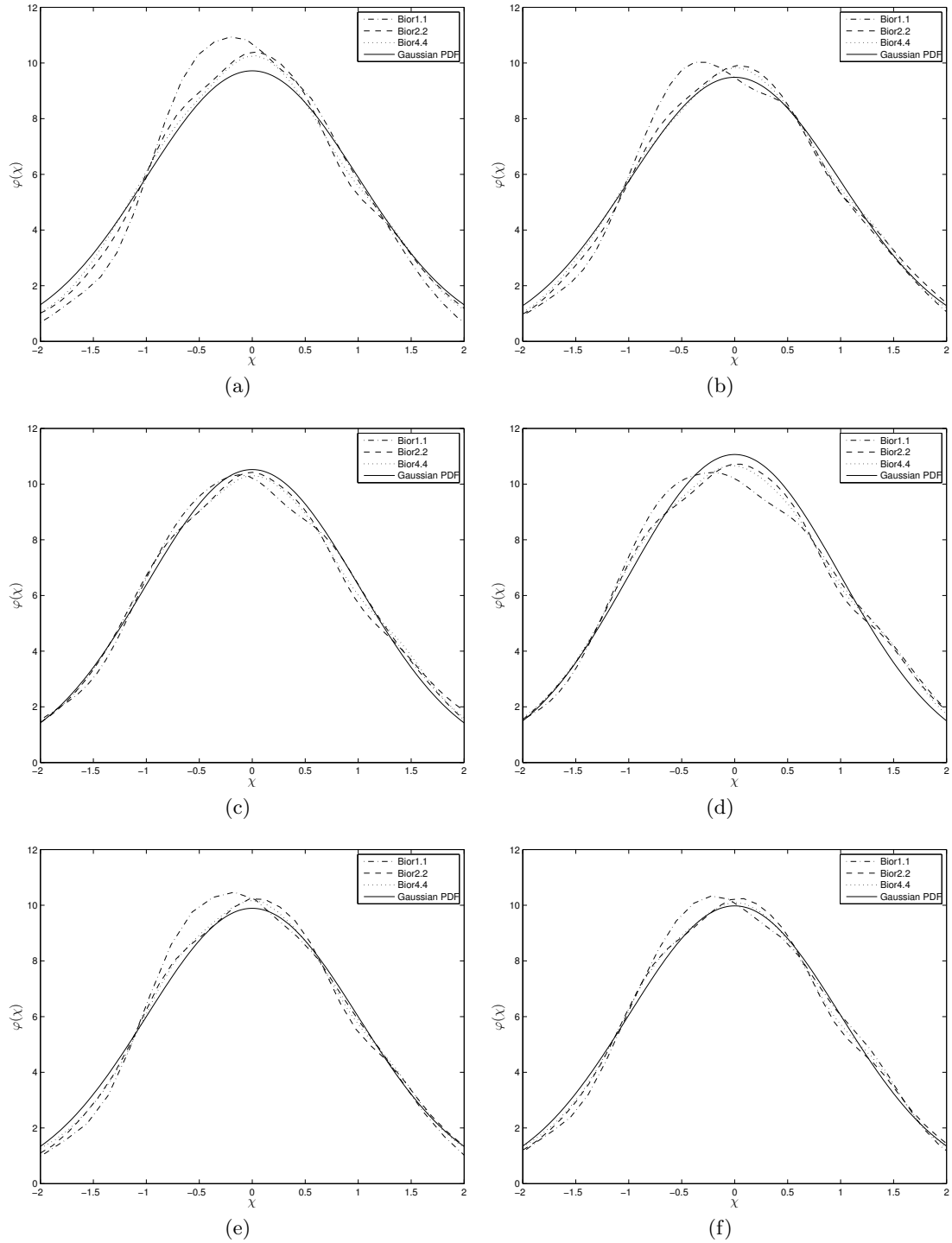


Figure 4.2: Distribution of wavelet coefficients at the first decomposition level, for signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f). Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal

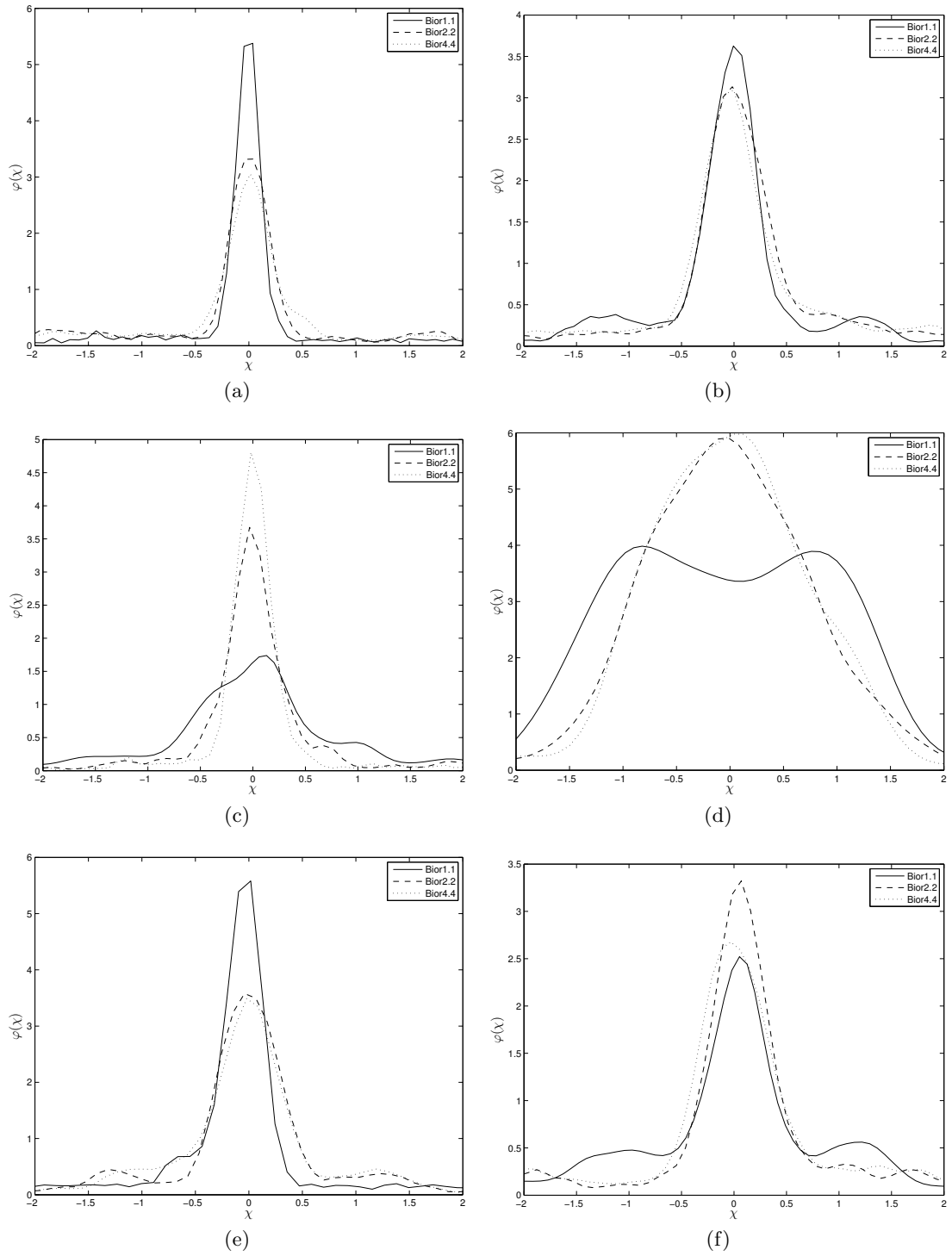


Figure 4.3: Distribution of wavelet coefficients at the last decomposition level, for signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f). Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal

reconstructed using longer and smoother wavelet basis. The shorter basis and, in the extreme case, the *Haar* wavelet, do not represent the optimal basis choice. In this case, the PDF for the *Bior4.4* wavelet basis is decisively narrower and with a higher peak than the other two basis. It is followed by the *Bior2.2* and finally the *Haar* wavelet.

The established connection between the suitability of a given wavelet basis for a specific signal and the distribution of wavelet coefficients at the last decomposition level, gives way to the assumption that the same connection should exist between the coefficients distribution and efficiency of the adaptive algorithm of the ICI-EPL denoising scheme. For correctly chosen parameter Γ value, the transform will adapt well to the local signal properties, and result in a higher quality reconstructed signal. The assumption is that the PDF of the wavelet coefficients at the last decomposition level takes narrower shape and have higher peak than in the case of poorly chosen Γ, for which the transform is unable to adapt well. Fundamentally, the distribution of wavelet coefficients at the last decomposition level provides a coarse measure of the quality of adaptation.

Modeling the Distribution of Wavelet Coefficients at the Last Decomposition Level

It was shown in Fig. 4.2 that PDFs of wavelet coefficients at the first decomposition level are very close to the Gaussian PDF. To describe the PDF at the last decomposition level, the Generalized Gaussian PDF will be used:

$$\begin{aligned} p(x) =& a\, exp(-|bx|^c), \\ a =& \frac{bc}{2\Gamma(\frac{1}{c})}, \\ b =& \frac{1}{\sigma_\chi}\sqrt{\frac{\Gamma(\frac{3}{c})}{\Gamma(\frac{1}{c})}}, \end{aligned} \tag{4.4}$$

where $\Gamma()$ is the Gamma function (Abramowitz and Stegun, 1964) and should not be confused with the ICI parameter Γ. Parameter c is the distributions shape parameter, for which there are two special cases. For $c = 2.0$, the GGD takes shape of the Gaussian PDF, while for the $c = 1.0$, the GGD takes shape of the Laplacian PDF. As the shape parameter is further lowered, the resulting GGD becomes narrower, while its peak gets higher, Fig. 4.4.

As already noted, wavelet coefficients distribution is neither unique nor constant, so there is no single value of the shape parameter c which would provide the optimal model for any given actual PDF. Instead, different parameter values are used in different cases and usage scenarios.

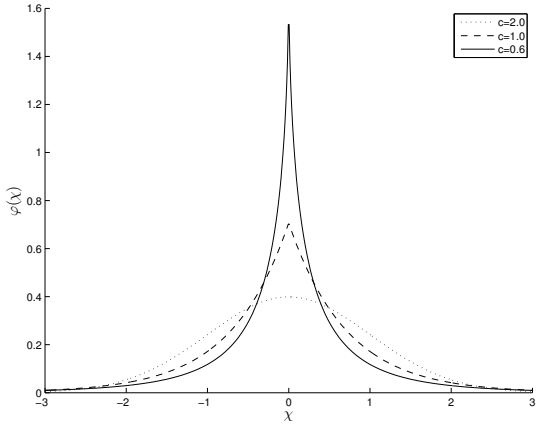


Figure 4.4: Generalized Gaussian distribution for different values of the shape parameter c

4.1.2 Influence of Noise

Noise Level

At low noise levels, wavelet transforms can very efficiently remove noise from a given noisy signal. If appropriate wavelet basis is chosen, even very fine original signal details can be reconstructed. As the noise level rises, the original signal properties can become substantially masked by the noise. Even if the chosen wavelet basis provides excellent fit to the local signal properties of a clean signal, the ambiguity introduced by the noise, results in a lower transform performance. Should our hypothesis stand, it will also result in a distribution of wavelet coefficients which can be modeled by a GGD with progressively larger shape parameter values, as the noise level rise. This is true for all test signals, as shown in Fig. 4.5. Again, the most appropriate of the three wavelet basis was chosen for each signal. Gaussian noise with standard deviation $\sigma_\epsilon \in \{5\%, 10\%, 15\%\}$ of total signal magnitude was added to the test signals. As there are many factors impacting the coefficients distribution, influence of noise level differs between the signals. Still, it is evident that increased noise levels lead to a wavelet coefficients distribution which can be modeled using a GGD with higher shape parameter values.

Noise Type

To predict influence of noise type on distribution of wavelet coefficients at the last decomposition level, we should briefly analyze $\mathcal{W}(\epsilon(x))$, i.e., wavelet transform of the noise component of $z(x)$. In this case, input signal to the transform is pure noise, and each signal sample can be considered to be a realization of an independent and identically distributed random variable (IID). Both, the coarse signal approximation and the wavelet coefficients are simple weighted sums of the IIDs. According to the central limit theorem, whichever was the initial noise distribution, given adequate sample size, the sum of the IIDs will always converge to a Gaussian distribution. Therefore, we do not expect the noise

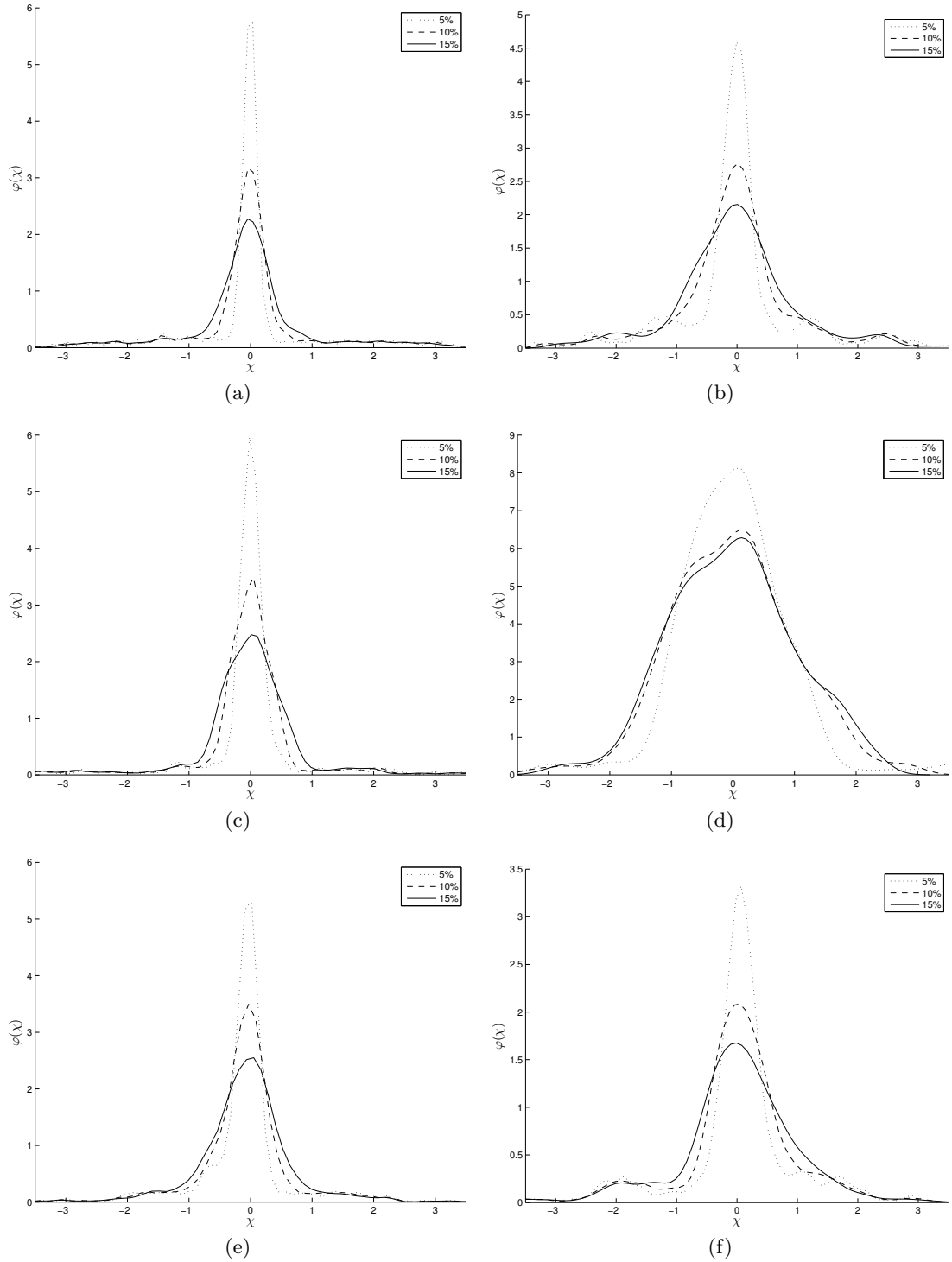


Figure 4.5: Distribution of wavelet coefficients at the last decomposition level for signals with additive Gaussian noise with standard deviation $\sigma_\epsilon \in \{5\%, 10\%, 15\%\}$ of the total signal magnitude. Signals used are: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

type to have significant impact on the coefficients distribution at the last decomposition level. This can be confirmed by analyzing Fig. 4.6. Gaussian, Poisson and Laplacian noise, with $\sigma_\epsilon = 10\%$ of total signal magnitude were added to the test signals. It can be seen that, in general, the PDFs share very similar shape, although the Gaussian noise corruption tends to result in a slightly wider PDF curve.

4.1.3 Number of Decomposition Levels

In the foregoing discussions we were repeatedly referring to the wavelet transforms last decomposition level, without deciding how many levels are actually to be used. Optimal number of decomposition levels varies, based on a signal and noise content of $z(x)$. Larger than optimal number of levels do not bring further improvements in denoising efficiency or even lowers it slightly, but also incur high cost in terms of unnecessary computational complexity. Since our discrete time test signals are considered to be samples of their respective continuous time functions, sampled at different sampling frequencies, number of decomposition levels used will be a function of signal length, as per the Tab. 4.1.

Table 4.1: Number of decomposition levels used, based on the test signal length

Signal length	256	512	1024	2048	4096
Total decomposition levels	3	4	5	6	7

Higher number of wavelet decomposition levels, for signals sampled at higher sampling frequencies, allow for finer separation of noise and original signal properties and a more compact signal representation in the transform domain. In line with that, we expect that the wavelet coefficient PDFs will be narrower and with a higher peak, than in the case of test signals sampled at lower frequencies. The effect is shown in Fig. 4.7.

4.2 Considering Magnitudes of Wavelet Coefficients

In the lifting scheme realization of wavelet transforms (Fig. 2.1), wavelet coefficients represent prediction errors, i.e., the difference between the real values of samples from one signal phase, and prediction of their values, based on the samples from the other signal phase. It is possible to evaluate the performance of the adaptation algorithm, by analyzing magnitudes of the coefficients. In general, if the ICI algorithm adapts the predictor to the signal well enough, the prediction errors are expected to stay small. The small errors are assumed to account for noise and allow for efficient signal denoising using the wavelet thresholding. The fact was used in (Claypoole, Baraniuk, and Nowak, 1998). Claypoole et al. proposed an adaptive algorithm in which wavelet basis resulting in the smallest prediction error was chosen at each signal point.

In case of the ICI-EPL there is an additional factor to consider. As the transform tries to avoid edges in a signal, not only that most of the prediction errors should stay small, but, since many edges are retained in the coarse signal approximation, there should

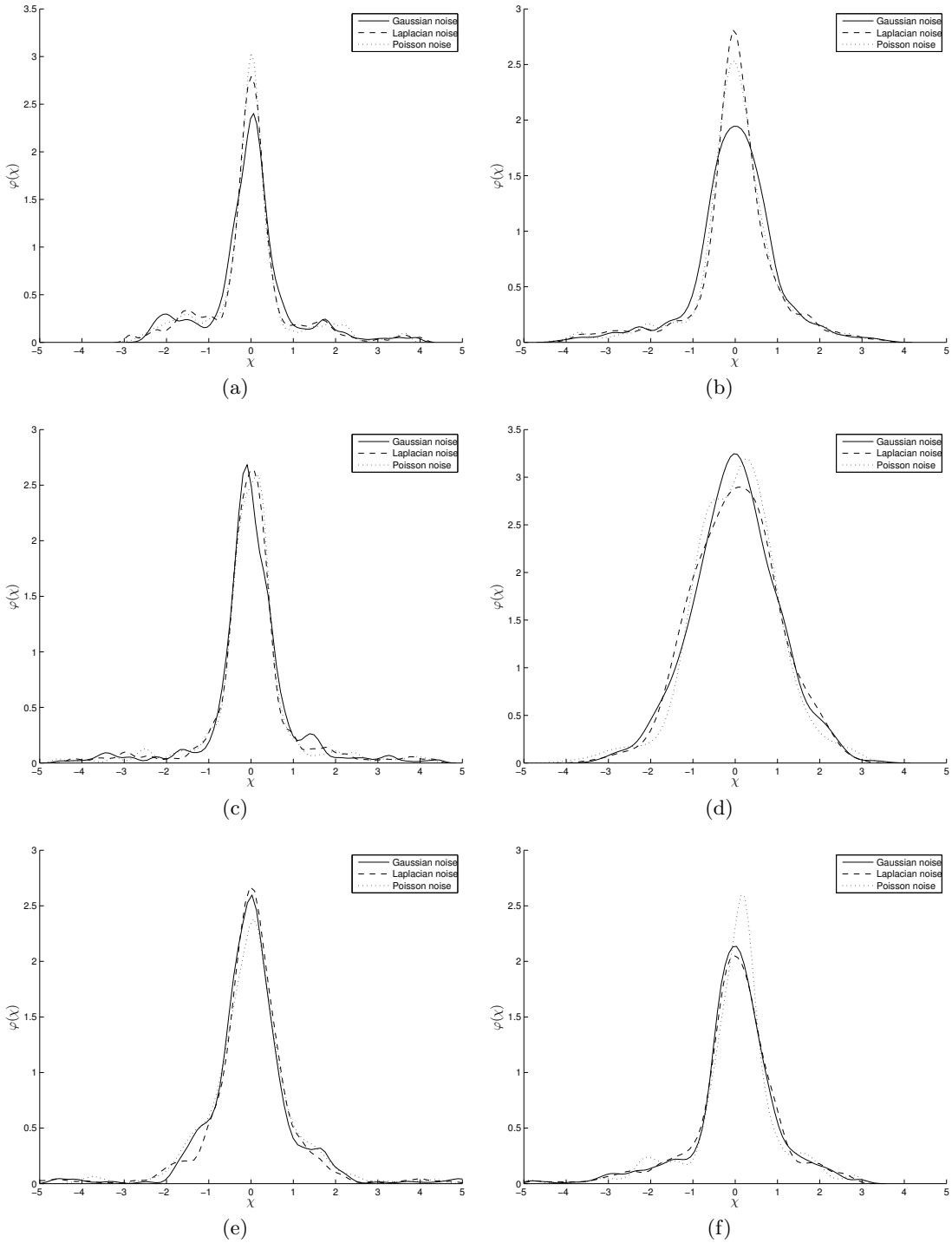


Figure 4.6: Distribution of wavelet coefficients at the last decomposition level, for signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f). PDFs for Gaussian, Poisson and Laplacian additive noise with $\sigma_\epsilon = 10\%$ of total signal magnitude are shown

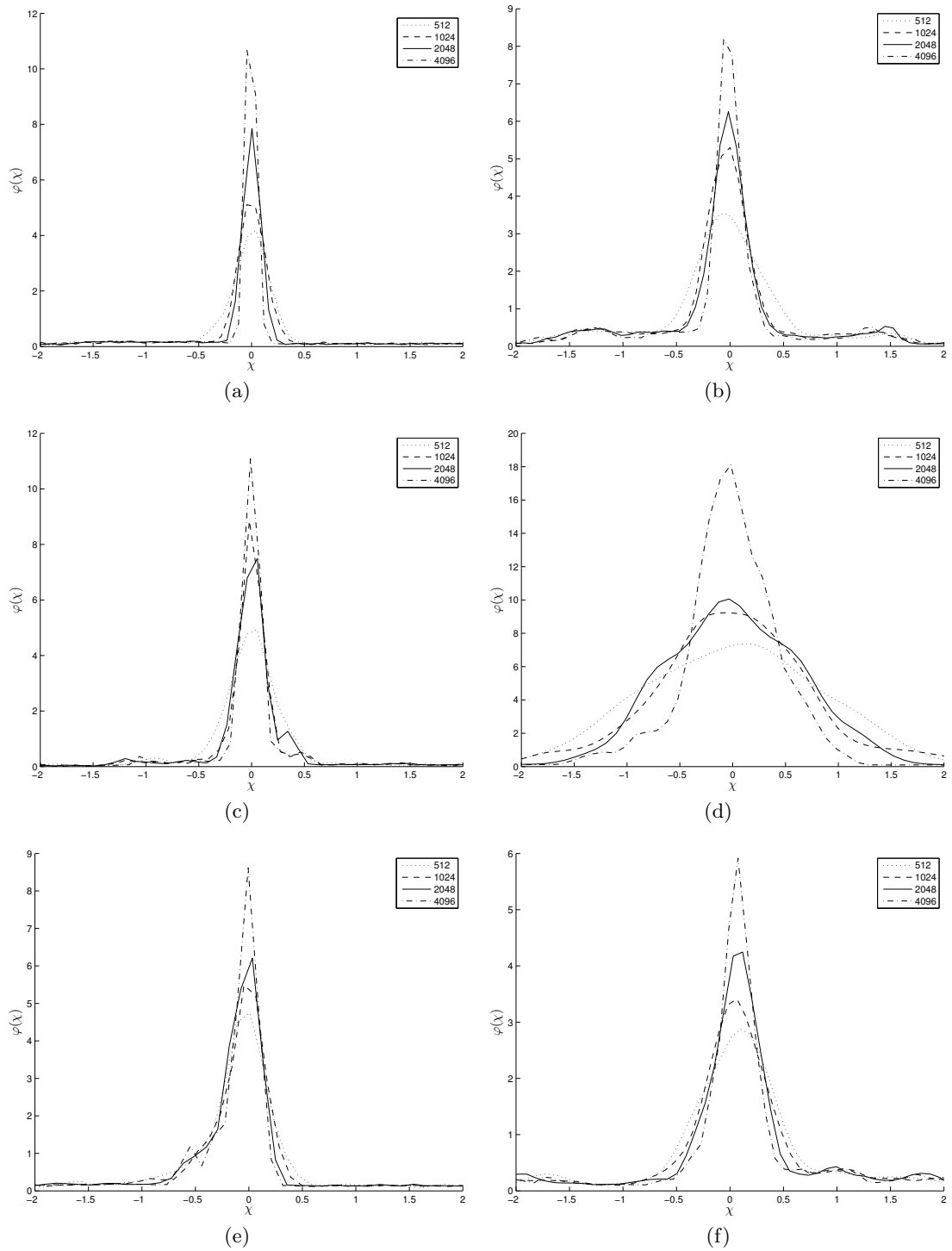


Figure 4.7: Distribution of wavelet coefficients at the last decomposition level, for various signal lengths: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f). Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal

also exist minimal number of very large wavelet coefficients. The difference between the conventional wavelet basis and optimal/sub-optimal ICI-EPL is shown in Fig. 4.8.

The figure shows wavelet coefficients at the last decomposition level for the *Blocks* signal, when using the *Haar* and the ICI-EPL wavelet transforms. The coefficients for the *Haar* wavelet transform contain a number of very large coefficients which are missing in the ICI-EPL case with optimal Γ parameter value (Γ_o). Even with a poorly chosen Γ value (Fig. 4.8 (c) and (d)), number of large coefficients is still smaller than in the *Haar* case.

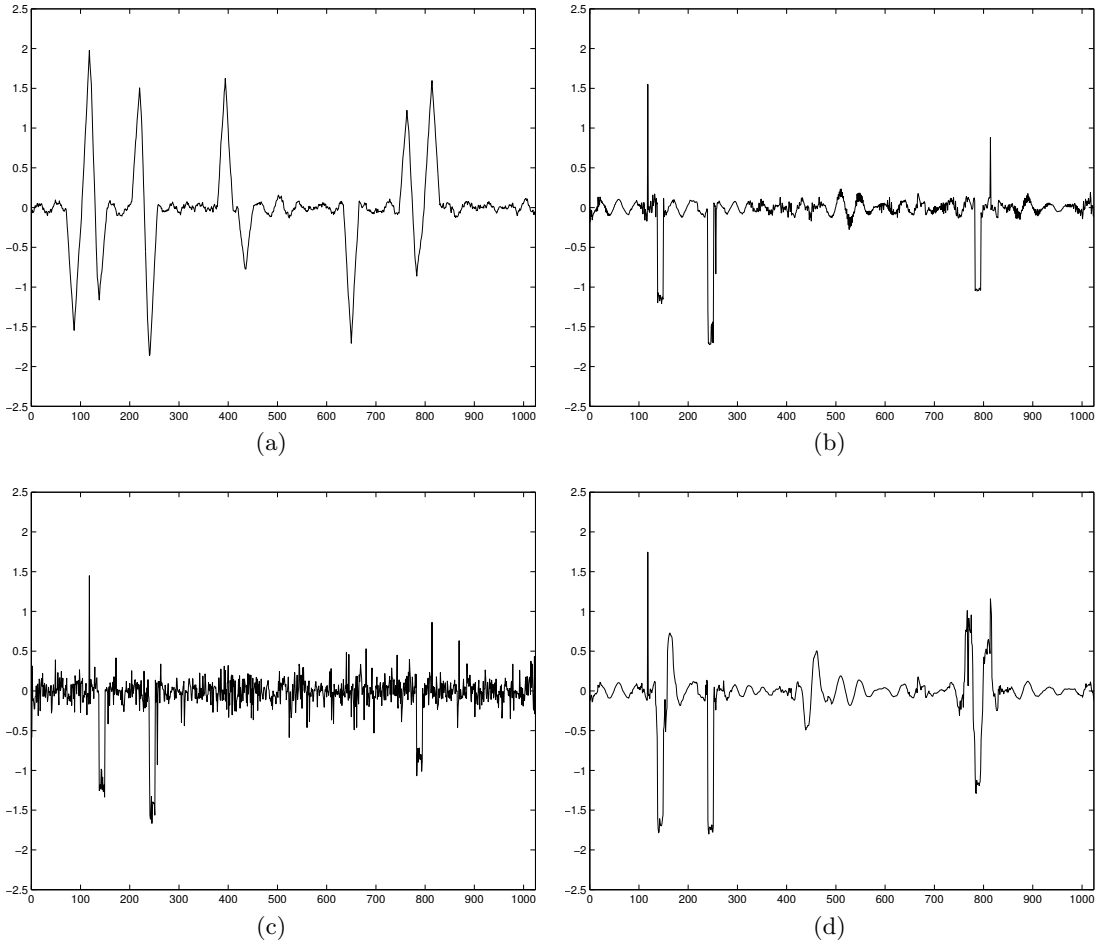


Figure 4.8: Wavelet coefficients at the last decomposition level for *Blocks* signal. *Haar* wavelet (a), ICI-EPL with optimal Γ parameter value, Γ_o (b), ICI-EPL with $\Gamma < \Gamma_o$ (c) and ICI-EPL with $\Gamma > \Gamma_o$ (d)

As a measure of adaptivity performance, that is based on the existence of very large wavelet coefficients, a simple sum of squared errors (i.e. wavelet coefficients) at the last decomposition level can be used:

$$SSE_\Gamma = \sum d_N(\Gamma)^2, \tag{4.5}$$

where N is the total number of decomposition levels. We expect that the transform which better adapts to a signal will yield a lower SSE_Γ values, than the one with less efficient adaptation.

Fig. 4.9 shows normalized values of the SSE_Γ measure across a range of the parameter Γ values. Optimal Γ value, Γ_o, is depicted by a vertical dotted line. Signal length is 1024 samples and additive Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal. Measure was taken at the last decomposition level. We can see that for the *Blocks*, *HeaviSine* and *Piece-polynomial* signals, the measure has minimum close to the Γ_o. For the *Bumps* signal, the result is still reasonable but in case of the *Doppler* or the *Piece-regular* signals, the Γ values suggested by the measure are unacceptable.

Fig. 4.10 shows normalized values for the same signal and length, but with the additive Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. In this case, results are much better and more consistent.

Simulations showed that, barring the class of a signal, the measure efficiency roughly depends on two factors: the noise level and the signal length. Generally, the efficiency improves as the noise level or signal length increases, while it declines for low noise levels or signals sampled at lower frequencies.

The SSE_Γ measure showed definite connection with the efficiency of the adaptive algorithm. In some cases it came close to pinpointing the optimal Γ parameter value but, unfortunately, it does not provide consistent and reliable enough results across the range of signal classes, lengths and noise levels. As such, it cannot be used as a standalone method for the parameter Γ value selection. However, it can be used to improve the efficiency of the method based on the distribution of wavelet coefficients, which will be proposed in the next section.

4.3 Selecting the Γ Parameter

We made a hypothesis, establishing a link between the distribution of wavelet coefficients at the last decomposition level, and the suitability of a given wavelet basis for the signal to be analyzed. After thorough examination of the coefficients distribution, and changes in the distribution caused by each of the influencing factors, we now consider the hypothesis to be valid. To summarize, the distribution of wavelet coefficients which is narrower and has a higher peak, points to a wavelet basis that better fits the signal. Such distributions can be modeled with GGDs with smaller shape parameter values, than in the case of choosing less appropriate wavelet basis.

The findings allow us to devise and formalize two methods for automated selection of the ICI-EPL parameter Γ. One method decides the parameter value by simple comparison of wavelet coefficient distributions peakedness, as quantified by the kurtosis measure. The other one compares the empirical distribution of wavelet coefficients at the last decomposition level to a hypothetical GGD the coefficients would come close to taking,

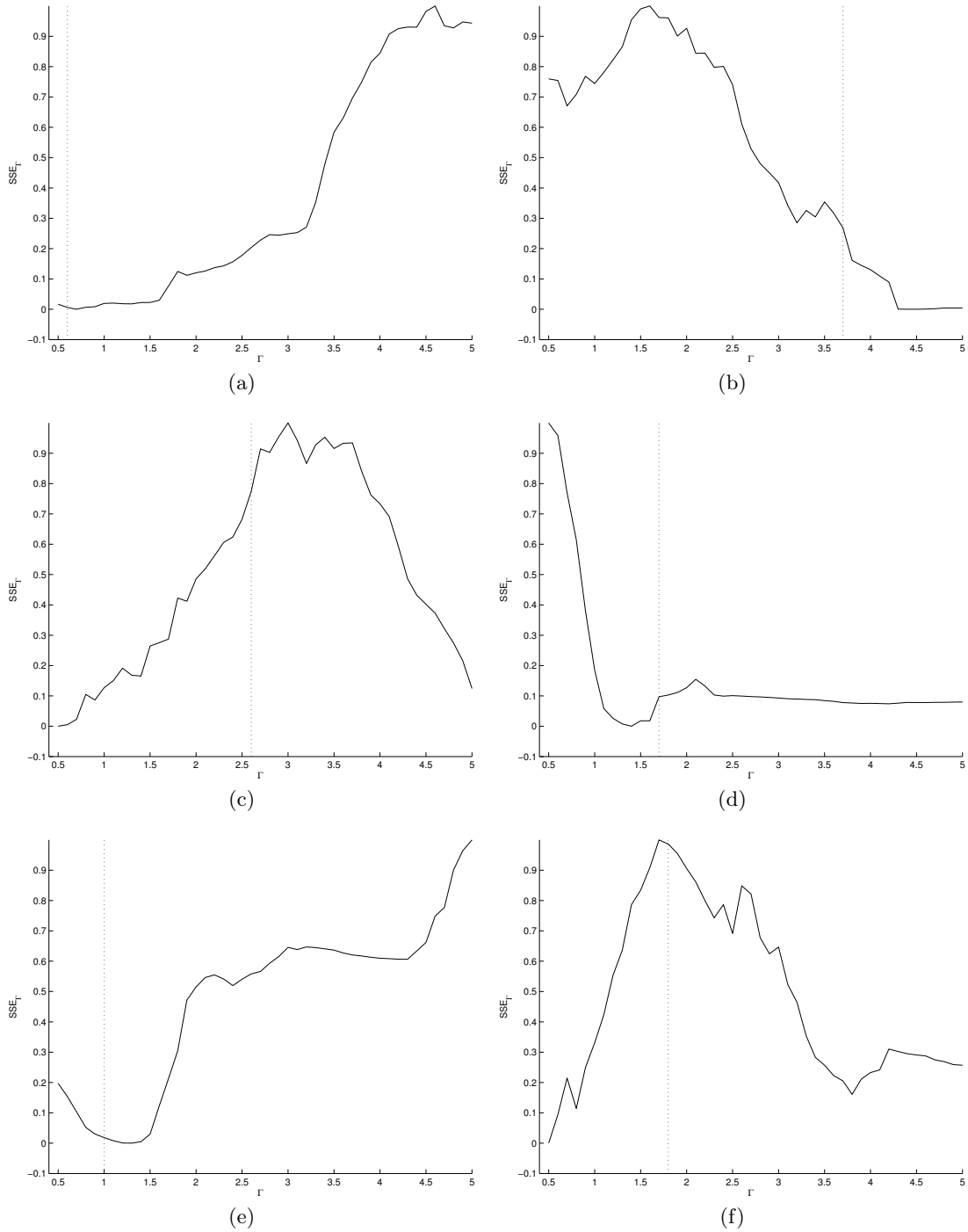


Figure 4.9: SSE_Γ of wavelet coefficients at the last decomposition level, for: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f) signals. Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal

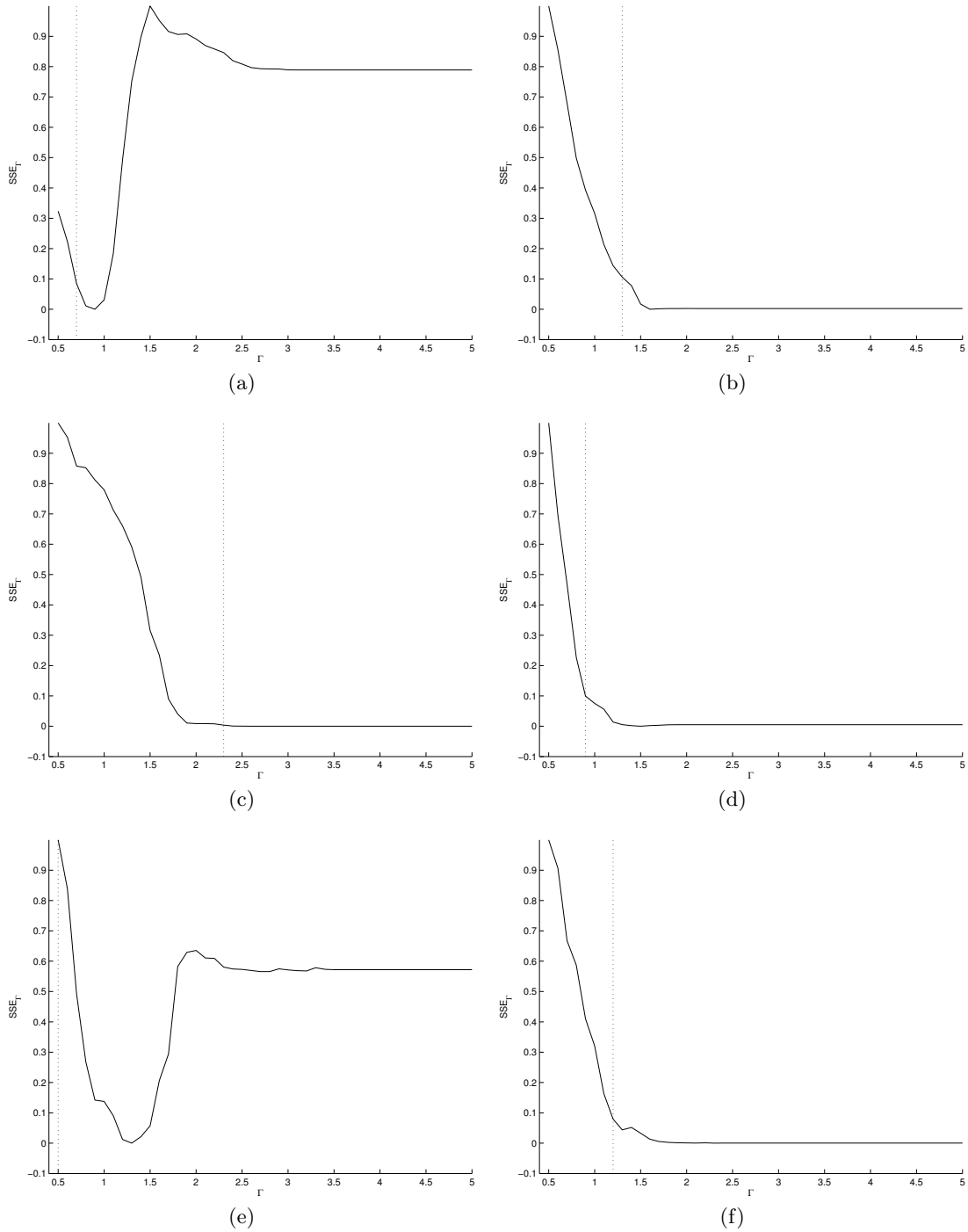


Figure 4.10: SSE_Γ of wavelet coefficients at the last decomposition level, for: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f) signals. Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude was added to each signal

in case there was a near-perfect fit of a wavelet basis to a given signal.

4.3.1 Comparing Distribution Peakedness

To quantify the peakedness of a distribution, a kurtosis measure is used:

$$k = \frac{E(x-\mu)^4}{\sigma^4}, \tag{4.6}$$

where μ is the mean of x, σ is the standard deviation of x, and $E(t)$ represents the expected value of the quantity t. In line with the previous discussions, we expect the wavelet coefficient distributions for correctly chosen Γ values to be more leptokurtic than in the case of sub-optimal parameter Γ values.

Distribution kurtosis provides us with information on compactness of noisy signal, $z(x)$, representation in the transform domain. Although influence of noise on coefficients distribution is diminishing, as we go towards higher decomposition levels, it still exists. If original signal, $y(x)$, was available, we could check the coefficients distribution kurtosis to see how well the transform adapts to the signal itself. Similar effect can be achieved by calculating kurtosis of thresholded wavelet coefficients . By choosing the threshold level of, for instance:

$$t = 3\sigma_\epsilon, \tag{4.7}$$

we can eliminate most of the noise influence. Inevitably, some of the original signal information is also going to be missing in the thresholded coefficients, but what remains should be enough to judge the adaptive algorithm efficiency.

The parameter Γ value selection method based on the distribution kurtosis is mostly straightforward. First, we must decompose a noisy signal, using the ICI-EPL. Decomposition has to be performed for a full range of reasonable parameter Γ values. Matrices of both, the original and the thresholded, wavelet coefficients distribution kurtosis are populated by calculating the kurtosis at each decomposition level and for each value of the parameter Γ:

$$\begin{aligned} \mathbf{k}_{(\mathbf{\Gamma},\ \mathbf{j})} &= k(d(\Gamma, j)), \\ \mathbf{k}_{\mathbf{t}(\mathbf{\Gamma},\ \mathbf{j})} &= k(d_t(\Gamma, j)), \end{aligned} \tag{4.8}$$

where $d(\Gamma, j)$ are wavelet coefficients at level j of the transform, using a specific parameter Γ value. Thresholded coefficients and respective distribution kurtosis are denoted by the index t.

By finding the maximum kurtosis, $max(\mathbf{k})$, we find a specific case for which a noisy signal can be most compactly represented in the transform domain. Noting the decomposition level at which the maximum is achieved, j_m, we can now analyze the $\mathbf{k_t}$ and conclude the Γ value selection method. The Γ value shall be selected as a value for which the $\mathbf{k_t}$, at level j_m, has a maximum, i.e., for which the original signal is most compactly

represented in the transform domain, at the level j_m. The value is denoted as Γ_k^*.

Most often, the $max(\mathbf{k})$ is found on levels $j >= 3$. However, in some cases, mostly for short signals, it can also be found at lower decomposition levels, $j \in \{1, 2\}$. Regardless of the original signal type, it is difficult to expect that its impact on the coefficients distribution at these low levels is large enough to represent a reliable measure of the adaptive algorithm performance. To improve the selection method efficiency in such scenarios, we can slightly modify it, such that the $max(\mathbf{k})$ is only searched for at decomposition levels $j >= 3$. We denote the Γ value obtained by using the modified selection method as Γ_{km}^*.

4.3.2 Comparing the Empirical and Hypothetical Distributions

Second method for the parameter Γ value selection is based on a modeling of wavelet coefficients distribution at the last decomposition level. Should there be a perfect fit of a chosen wavelet basis to a signal being decomposed, the distribution of wavelet coefficients at the last decomposition level would be very narrow and have a high peak. We can model it with a GGD with a small shape parameter c. Our strategy is to find a Γ value for which the empirical coefficients distribution is most similar to that hypothetical GGD. In order to compare the distributions, we first have to perform decomposition of a given signal, using the ICI-EPL transform. A full range of reasonable Γ parameter values is tried, and wavelet coefficient at the last decomposition level (N) are recorded:

$$d_N(\Gamma). \tag{4.9}$$

The empirical distribution of wavelet coefficients at the last decomposition level, denoted as $WCD(d_N(\Gamma))$, is tested for a goodness-of-fit against the appropriate hypothetical GGD:

$$t(WCD(d_N(\Gamma)), GGD(c, \sigma^2)), \tag{4.10}$$

where t is the test statistic. According to observations in the first section of this chapter, actual shape parameter c value, which defines the hypothetical GGD, is mostly influenced by a signal length. Appropriate values for each of the signal lengths (in effect, each of the signal sampling frequencies) are found empirically.

By analyzing the parameter Γ selection method performance graphs, shown in Appendix C, we can see that the optimal performance is achieved for a limited range of shape parameter values. Appropriate value is chosen as a value which, approximately, lies in the middle of a given range. Chosen shape parameter values, shown in Tab. 4.2, should work well for a variety of different signal classes. If only a specific class of signals is targeted, they could also be further adjusted or optimized.

To improve efficiency of the Γ parameter value selection, the test statistic is multiplied by the sum of squared wavelet coefficients at the last decomposition level (4.5), resulting

Table 4.2: Appropriate GGD shape parameter values, based on a signal length (sampling frequency)

Signal length	256	512	1024	2048	4096
GGD shape parameter value	0.5	0.4	0.35	0.3	0.25

in the final score for the given Γ value:

$$T(\Gamma) = t(WCD(d_N(\Gamma)), GGD(c, \sigma^2)) \cdot SSE_\Gamma. \tag{4.11}$$

The final value of the Γ parameter (Γ^*_{GGD}) shall be selected as the value for which the score function has a minimum:

$$T(\Gamma^*_{GGD}) = \min(T(\Gamma)). \tag{4.12}$$

In rare occasions, when the score function has minimum for more than one Γ parameter value, we will opt to choose the smallest value.

Testing the Goodness-of-Fit

There are many statistical tests for fitting of data to a distribution (Thode, 2002). Each of the tests has certain advantages and disadvantages. Neither of them can be considered to be the most powerful one, as each of the tests, in certain cases, outperforms the others. We will evaluate the Γ selection method performance when using three commonly known tests: the Kolmogorov-Smirnov test, the Pearson chi-squared test and the Cramer-von-Mises criterion.

The Kolmogorov-Smirnov (KS) test (Chakravarti, Laha, and Roy, 1967) provides a measure of equality of two continuous probability distributions. It is used to decide whether a sample comes from a population with a specific, reference, distribution. The KS statistic quantifies the difference between the actual distribution function of the sample, and the reference distributions cumulative distribution function (Thas and Ottoy, 2002). The null hypothesis is that the reference distributions fits the data, i.e., the sample is drawn from a reference distribution. The KS test statistic is defined as:

$$D = \max_{1<i<J} |F_n(i) - F(i)|, \tag{4.13}$$

where J is the sample length, F_n is the sample empirical distribution function (Kaplan-Meier estimate), and F is the cumulative distribution function of a reference distribution. In case the sample is drawn from the reference distribution, the D-statistic almost surely converges to 0. The D-statistic is compared to the critical values of the Kolmogorov distribution to either accept or reject the null hypothesis. In our case, we are not concerned if the null hypothesis is accepted or rejected (in fact, it would probably be rejected most of the time). We are only interested in the test statistics value. It will allow us to compare

how close each of the empirical wavelet coefficient distributions (originating from the ICI-EPL transforms, using different Γ values), are close to our hypothetical (reference) distribution. As it is not necessary for the proposed Γ selection method, critical values or their determination is not discussed further.

The Pearson chi-square (χ^2) test (Snedecor and Cochran, 1967),(Thas and Ottoy, 2002) is another well known goodness-of-fit test. Unlike the KS test, the χ^2 test works on discrete distributions. This should not be considered as a limitation of the test since the continuous data can also be analyzed. Prior to using the test, it is necessary to bin the data, for instance, by calculating a histogram. Although not a limitation, it could be considered as a disadvantage, since the test statistic can change significantly, based on the way the data is binned. The null hypothesis is that the observed probability distribution do not differ from a specific theoretical distribution. To test the hypothesis, the data is divided into M bins, and the χ^2 test statistic is calculated:

$$\chi^2 = \sum_{i=1}^{M} \frac{(O_i - E_i)^2}{E_i}, \tag{4.14}$$

where O_i is the observed frequency for the bin i and E_i is a theoretical frequency asserted by the null hypothesis. The statistic can also be written using the probability density functions, resulting in the following form:

$$\chi^2 = M \sum_{i=1}^{M} \frac{(f_n(x_i) - f(x_i))^2}{f(x_i)}, \tag{4.15}$$

where f_n and f are, respectively, probability density functions of the sample data and of the theoretical distribution asserted by the null hypothesis. There is no optimal number of bins the data should be divided into, however, the result should be similar for any reasonable number of bins. One common approach is to tie the bin width to the sample standard deviation. This way, each bin, with the exception of the lower and upper ones, has the same width: $a\sigma$. The exact value for the a has to be chosen based on a distribution of a given sample.

The last test to be considered is the Cramer-von-Mises (CM) test. It is not as well known as the KS or the χ^2 tests but is often cited as an alternative to the KS test. It is used to judge how well the empirical distribution function fits to the reference theoretical cumulative distribution function. Again, the null hypothesis is that the sample comes from the reference theoretical distribution. The Cramer-von-Mises ω^2 criterion is defined as:

$$\omega^2 = \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x), \tag{4.16}$$

where $F_n(x)$ is the empirical distribution function of the sample and $F(x)$ is the reference

distribution function. From the criterion, follows the CM test (Anderson, 1962):

$$T = J\omega^2 = \frac{1}{12J} + \sum_{i=1}^{J}\left[\frac{2i-1}{2J} - F(x_i)\right]^2, \tag{4.17}$$

where J is the sample size, and x_i are observed values, sorted in increasing order.

In our case, we will be testing how well the wavelet coefficients at the last decomposition level fit the generalized Gaussian distribution (GGD) with a chosen shape parameter c. So, in each of the tests, the reference distribution will be the GGD and the empirical distribution will be the actual coefficients distribution. After substitution, the final expressions for the three test statistics are given by:

$$\begin{aligned} D_{\Gamma,c} &= \max_{1<i<J} |F_{d_N(\Gamma)} - F_{GGD_c}|, \\ \chi^2_{\Gamma,c} &= M\sum_{i=1}^{M}\frac{(f_{d_N(\Gamma)}(x_i) - f_{GGD_c}(x_i))^2}{f_{GGD_c}(x_i)}, \\ T_{\Gamma,c} &= J\omega^2 = \frac{1}{12J} + \sum_{i=1}^{J}\left[\frac{2i-1}{2J} - F_{GGD_c}(x_i)\right]^2, \end{aligned} \tag{4.18}$$

where $F_{d_N(\Gamma)}$ and $f_{d_N(\Gamma)}$ are, respectively, empirical distribution function and the probability density function of the wavelet coefficients at the last decomposition level, while F_{GGD_c} and f_{GGD_c} are, respectively, the cumulative distribution function and the probability density function of a GGD with the shape parameter c.

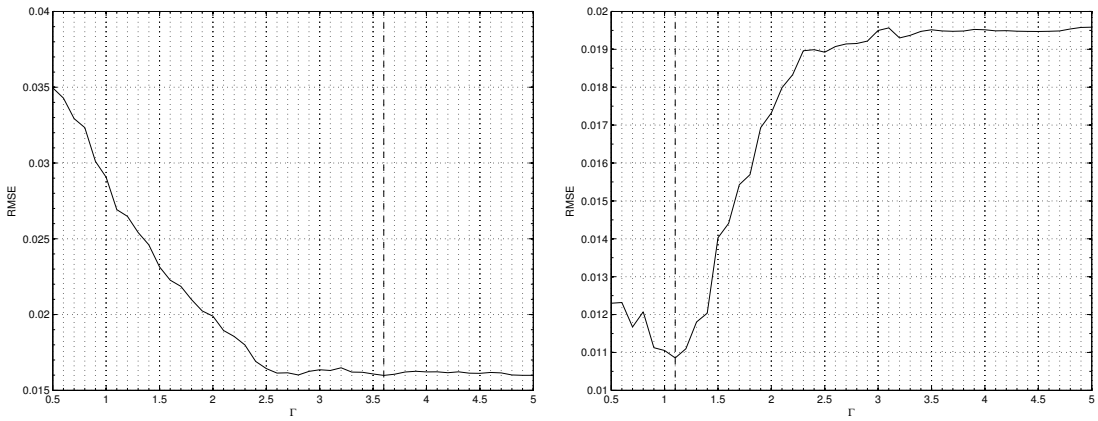
Although the tests are, generally, used to decide if a given sample is drawn from the reference distribution, this is not the case in our Γ selection method. As already noted, we do not need confirmation that the sample is drawn from a chosen GGD. We are only concerned of how close the actual distribution comes to the GGD. The closer it comes to, the more efficient the adaptation algorithm is expected to be. The three tests allow for the quantification of the "closeness".

4.4 Results and Discussion

To investigate efficiency of the proposed methods, and decide whether they can be considered to be generally applicable, we carry out multiple simulations. All three major factors directly influencing the method outcome are taken into account: the signal properties, the signal length and the level of additive Gaussian noise. To test the influence of signal properties, 6 different classes of signals are used – the same 1-D signals considered throughout the chapter: *Blocks*, *Bumps*, *Doppler*, *HeaviSine*, *Piece-polynomial* and *Piece-Regular*. Signal lengths (corresponding to different sampling frequencies), are $\{256, 512, 1024, 2048, 4096\}$. The effect of noise level is inspected by superimposing the white Gaussian noise, with standard deviation $\sigma_\epsilon \in \{0.05, 0.10, 0.15, 0.20\}$ of total signal magnitude. Values of the Γ parameter are in the $[0.5, 5]$ range, with increments of 0.1. In

order to remove the influence of particularities of a specific noise realization, experiments were performed with 10 different realizations and results were averaged.

Although both proposed methods try to determine the optimal Γ parameter value, their efficiency, cannot be evaluated by examining how close the selected value comes to the Γ_o. It is easy to comprehend this by looking at Fig. 4.11, in which root mean square errors (RMSE) of two denoised signals are plotted against the Γ values. In case of the *Doppler* signal (Fig. 4.11 (a)), the optimal Γ parameter value is 3.6 (as denoted by the vertical dashed line). However, the $\Gamma - RMSE$ curve shows clearly that almost the same denoising efficiency is achieved for a very wide range of the Γ values – $[2.6, 5]$. For instance, had the proposed method selected 5 as the Γ value it would have missed the optimal value by more than a half of the total range of Γ values. Regardless of the seemingly very poor value selection, the final result is an almost optimal denoising performance. On the other hand, in case of the *Piece-Polynomial* signal (Fig. 4.11 (b)), the $\Gamma - RMSE$ curve has a distinct minimum at $\Gamma = 1.1$. Had the proposed method selected 1.5 for the Γ value, it would have missed the Γ_o for only 4 increments. If one was to evaluate the method efficiency based on the closeness of the selected Γ parameter value to the actual optimal value, the conclusion would be drawn that the method performed better in the latter case. It would, obviously, be wrong conclusion as the denoising performance would have been degraded by $\approx 27\%$.



(a) (b)

Figure 4.11: Dependency of RMSE of denoised signals on the Γ parameter value, for: *Doppler* (a) and *Piece-Polynomial* (b) signals. Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal. Optimal Γ value is denoted by the vertical dashed line

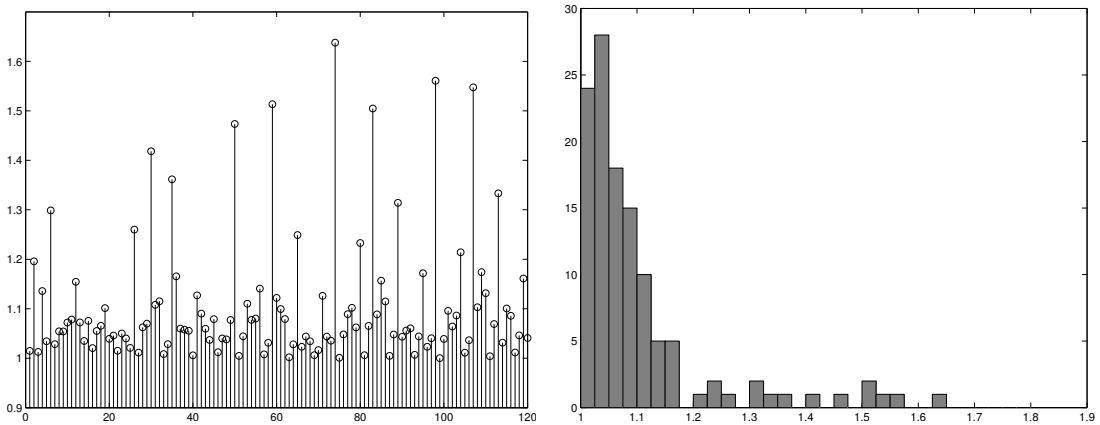
The closeness of the selected Γ parameter value to the optimal value could be interesting for gaining a very rough idea of the method performance, but for the correct evaluation of its efficiency, we have to consider the final result, i.e. the denoised signals. Quality of denoised signals is expressed in the terms of the root mean square error (RMSE). The ultimate goal is to select the exact optimal Γ parameter value, resulting in denoised signal

with the smallest RMSE. Unfortunately, it will most often not be the case. To evaluate the Γ selection method efficiency, we have to compare the $RMSE_{\Gamma^*}$ to a $RMSE_{\Gamma_o}$.

Tables containing RMSE values for all the test cases can be found in Appendix B.

4.4.1 Efficiency of the Kurtosis Based Selection Method

The efficiency of the kurtosis based parameter Γ value selection is depicted in Fig. 4.12. The $RMSE_{\Gamma^*_{km}}/RMSE_{\Gamma_o}$ ratios for each of the test cases are shown as points in Fig. 4.12 (a), while histogram of ratios is shown in Fig. 4.12 (b). We can see that in many of the test cases the automated Γ value selection performance is reasonable. In about half of the cases there is $< 5\%$ increase in RMSE and in about 70% of the test cases the increase is $< 10\%$. Unfortunately, in the remaining test scenarios there are many unacceptable results, which cause even more than a 50% increase in RMSE.



(a) (b)

Figure 4.12: $RMSE_{\Gamma^*_{km}}/RMSE_{\Gamma_o}$ ratios for each of the test cases (a) and histogram of ratios (b)

Results in Fig. 4.12 assume that the $max(\mathbf{k})$ is searched for on decomposition levels $j >= 3$. In Fig. 4.13, for reference only, we illustrate small improvements achieved by this minor modification. Only a subset of test cases, representing shorter signals, is shown, as there is no difference in method efficiency for longer signals.

4.4.2 Efficiency of the Distribution Model Based Selection Method

Since this is much more complex method, we will examine it more thoroughly. As in the previous case, the $RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o}$ ratio will be the main efficiency measure. The closer we get to 1 in a given test case, the more efficient the method is. Before going into greater details, let us first inspect the general improvements achieved by multiplying the goodness-of-fit test statistic with the SSE_Γ.

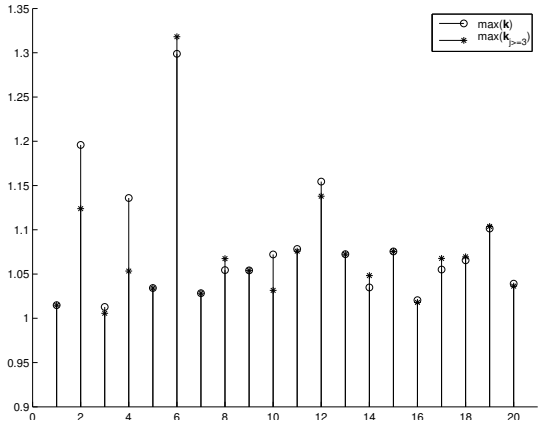


Figure 4.13: Improvements in selection method efficiency in case $max(\mathbf{k})$ is searched for on levels $j >= 3$

Multiplying the Test Statistic by the SSE_Γ

To illustrate the improvements, we show three score function plots. Fig. 4.14 shows plots of the KS test score function for the usual test signals. The function values are $RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o}$ ratios, in which Γ^*_{GGD} is found using the proposed, distribution model based, method for the parameter Γ value selection. Signal lengths are 1024 samples, and Gaussian white noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal. Fig. 4.15 shows the same type of plot, but this time the χ^2 test statistic was used. Signal length was also changed and it is now 256 samples. Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude was added to each signal. Finally, Fig. 4.16 shows the plot for the CM test statistic. Signal length is 4096 samples and Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude was added to each signal.

The figures were not shown to demonstrate or evaluate the efficiency of the selection algorithm, so the actual values of score functions are currently irrelevant. What is important, is the general difference between the test statistic score function and the SSE_Γ improved score function. Different goodness-of-fit test was used for each figure and a wide range of signal lengths and levels of additive Gaussian white noise were covered. It can be clearly seen that in all those situations the score functions improved by the SSE_Γ are consistently at or below the pure test statistic score functions. It confirms that the parameter Γ selection method makes better decision if the SSE_Γ is taken into account along with the distribution of wavelet coefficients.

The SSE_Γ is a poor criterion to be used as a complete Γ selection method, however, when used in conjunction with the findings about the distribution of wavelet coefficients at the last decomposition level, it successfully and significantly improves the, distribution based, selection methods efficiency and reliability. Even in the worst case scenario of application of the SSE_Γ measure, portrayed in Fig. 4.9 (c) and (f), if appropriate GGD shape parameter c is chosen, multiplication of the test statistic by the SSE_Γ will not have a detrimental effect to the overall efficiency of the selection method, as can be seen in Fig.

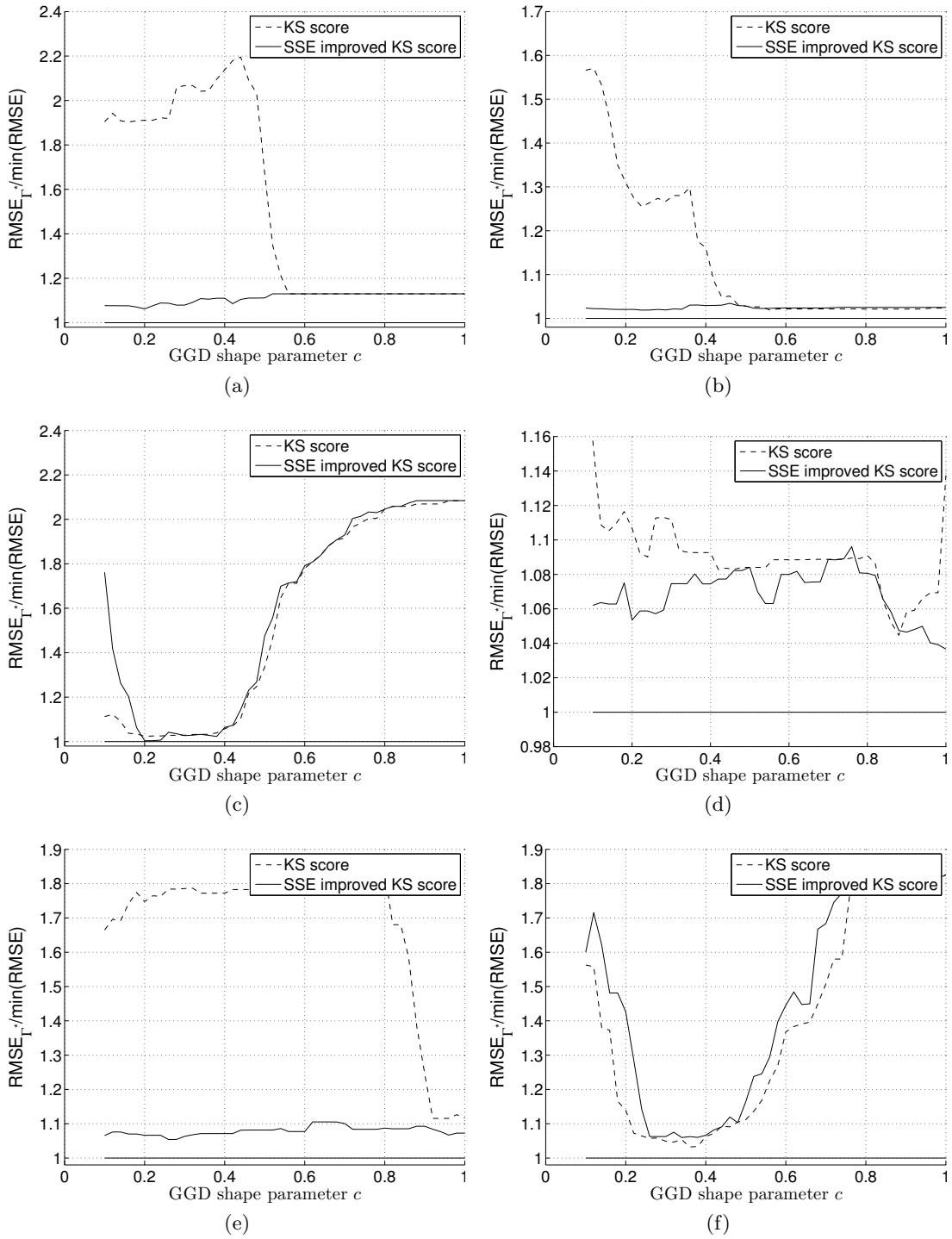


Figure 4.14: Score functions for the KS test statistic and the KS test statistic multiplied by the SSE_Γ, for: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f) signals. Signal length is 1024 samples. Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude was added to each signal

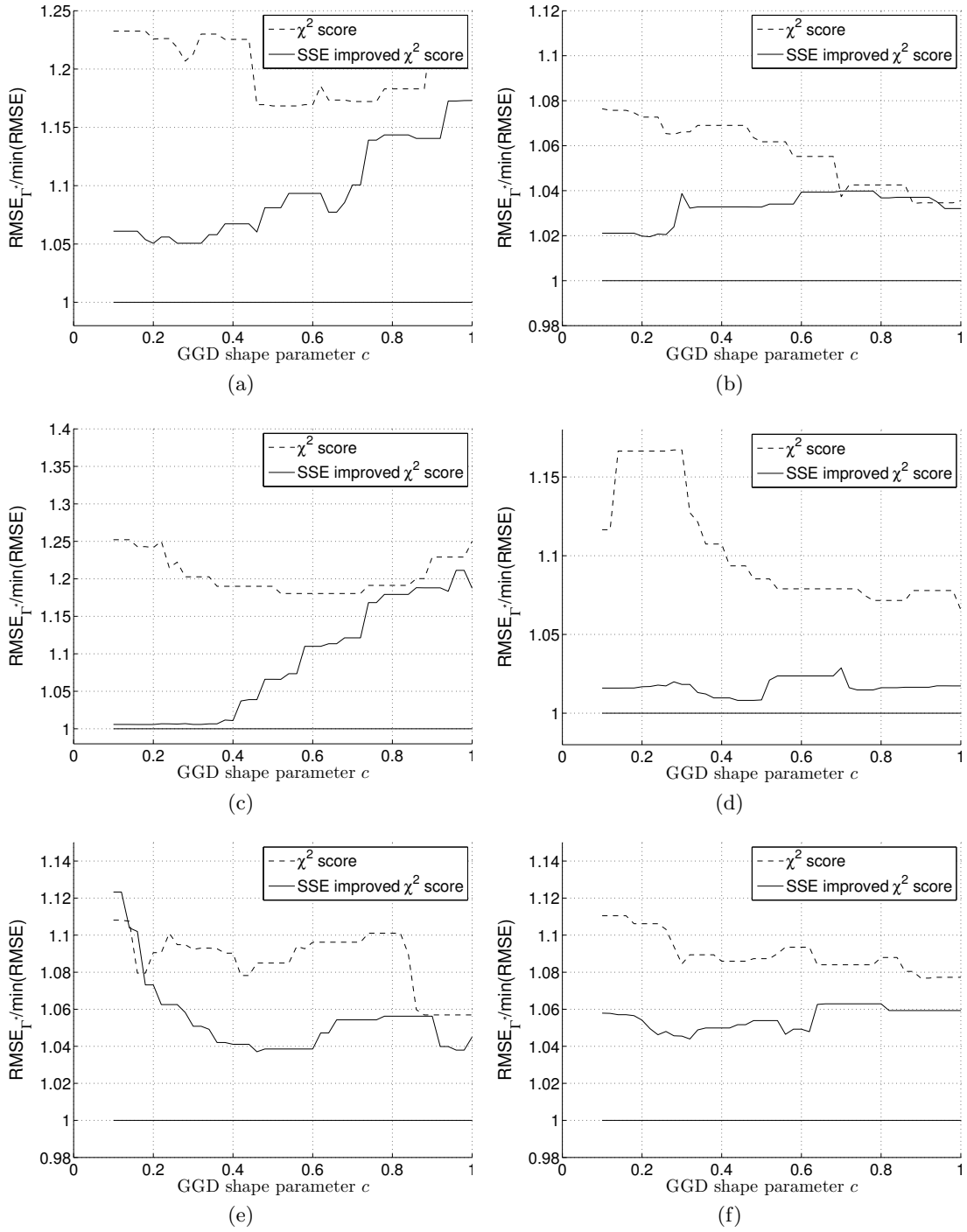


Figure 4.15: Score functions for the χ^2 test statistic and the χ^2 test statistic multiplied by the SSE_Γ, for: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f) signals. Signal length is 256 samples. Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude was added to each signal

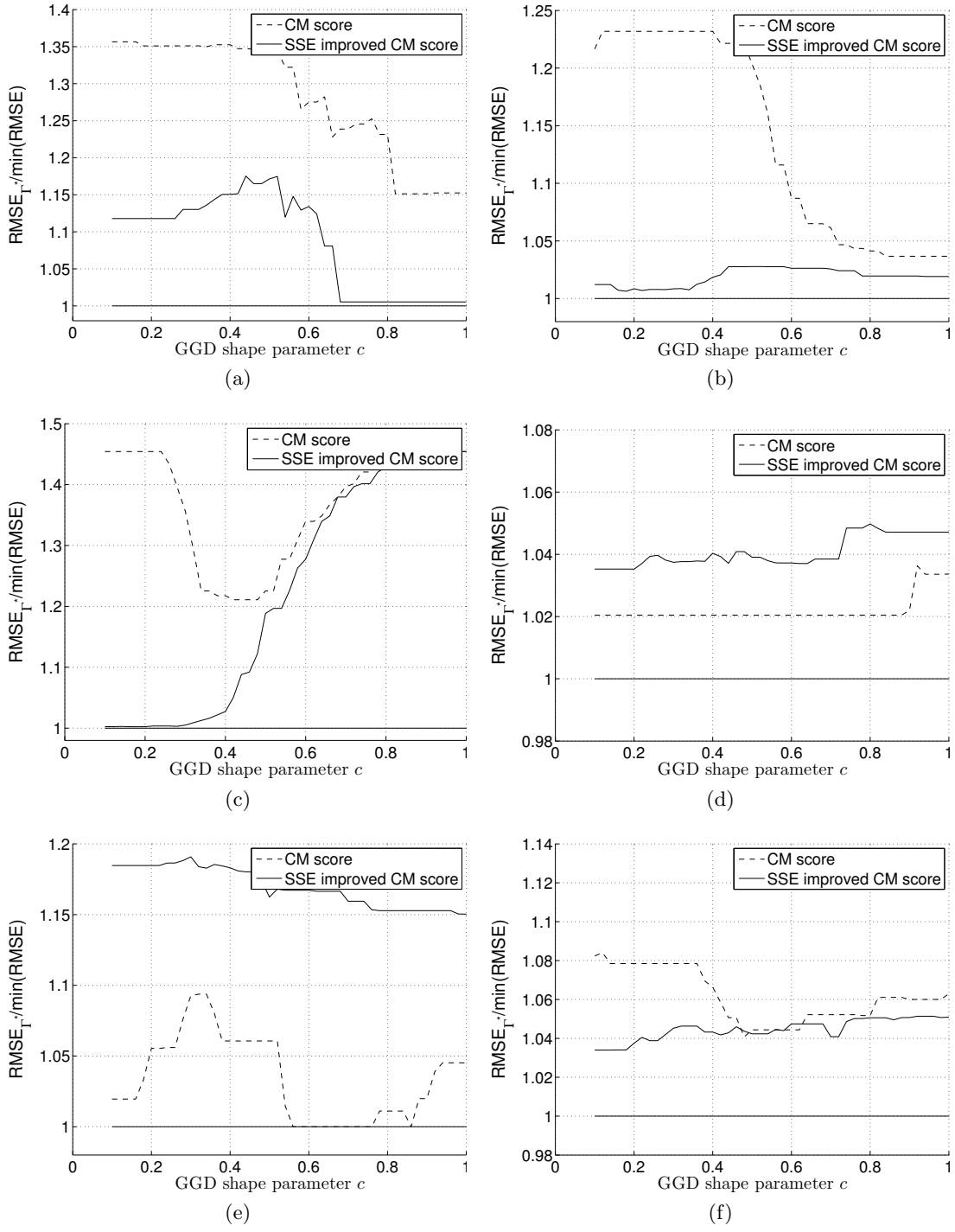


Figure 4.16: Score functions for the CM test statistic and the CM test statistic multiplied by the SSE_Γ, for: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f) signals. Signal length is 4096 samples. Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude was added to each signal

4.14 (c) and (f).

Comparing the Three Goodness-of-Fit Tests

As shown above, the efficiency of the method, is improved and made more stable and reliable by multiplying the test statistic by the SSE_Γ. We now compare the efficiency of the selection method for all three goodness-of-fit tests – the Kolmogorov-Smirnov (KS), the χ^2 (CS) and the Cramer-von-Mises criterion (CM). For the purpose, we show plots of the average $RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o}$ ratio achieved for a specific test, plotted against the full range of tested GGD shape parameter c values.

Only three plots are shown here for illustration purposes. They are chosen as an example of a near-best, an average and a near-worst case scenarios. The three could be considered to mostly represent the overall efficiency that can be expected from the proposed selection method. Plots for other signal lengths and noise levels tested, can be found in Appendix C.

Fig. 4.17 depicts efficiency of the selection method for 512 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Fig. 4.18 depicts the performance for signals with the same level of additive noise, but 4096 samples long. It can be seen that, for shorter signals, the method performed exceptionally well. On average, denoised *Blocks* and *Piece-Polynomial* signals, have less than 10% higher RMSE than for the optimal Γ value, while all the other test signals were below the 5% difference. The performance is similar in the case of longer signals, but larger discrepancies in performance of the *Blocks* and *Piece-Polynomial* signals, and the other signals can be observed. This is especially true for the *Piece-Polynomial* signal, for which, on average, achieved RMSE is more than 30% higher then the average optimal RMSE. The other signals are still bellow the 5% difference.

Although, in most cases, the performance for the *Blocks* and the *Piece-Polynomial* signal is lower than a performance for other signals, a different results can be found in the third example, Fig. 4.19. Depending on the goodness-of-fit test used, the average performance for the *Blocks* signal can get near the 5% mark, while the performance for the *Piece-Polynomial* signal is near the 5% mark, regardless of the test used. On the other hand, performance for the *Piece-Regular* signal is worse than in the previous two examples and comes, depending on the test being used between 10% and 20% worse than in the optimal case.

To compare performance of the proposed parameter Γ value selection method, based on the type of goodness-of-fit test being used, it is helpful to analyze the plots of the score functions, which can be found in Appendix C. To draw a conclusion on the preference of a particular test, two things have to be considered. One is the actual performance, i.e. the exact $RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o}$ values achieved. The other could be classified as a stability or a reliability. For instance, if both of the methods achieved $RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o} \approx 10\%$, the preferred method would be the one that shows similar performance for a larger range

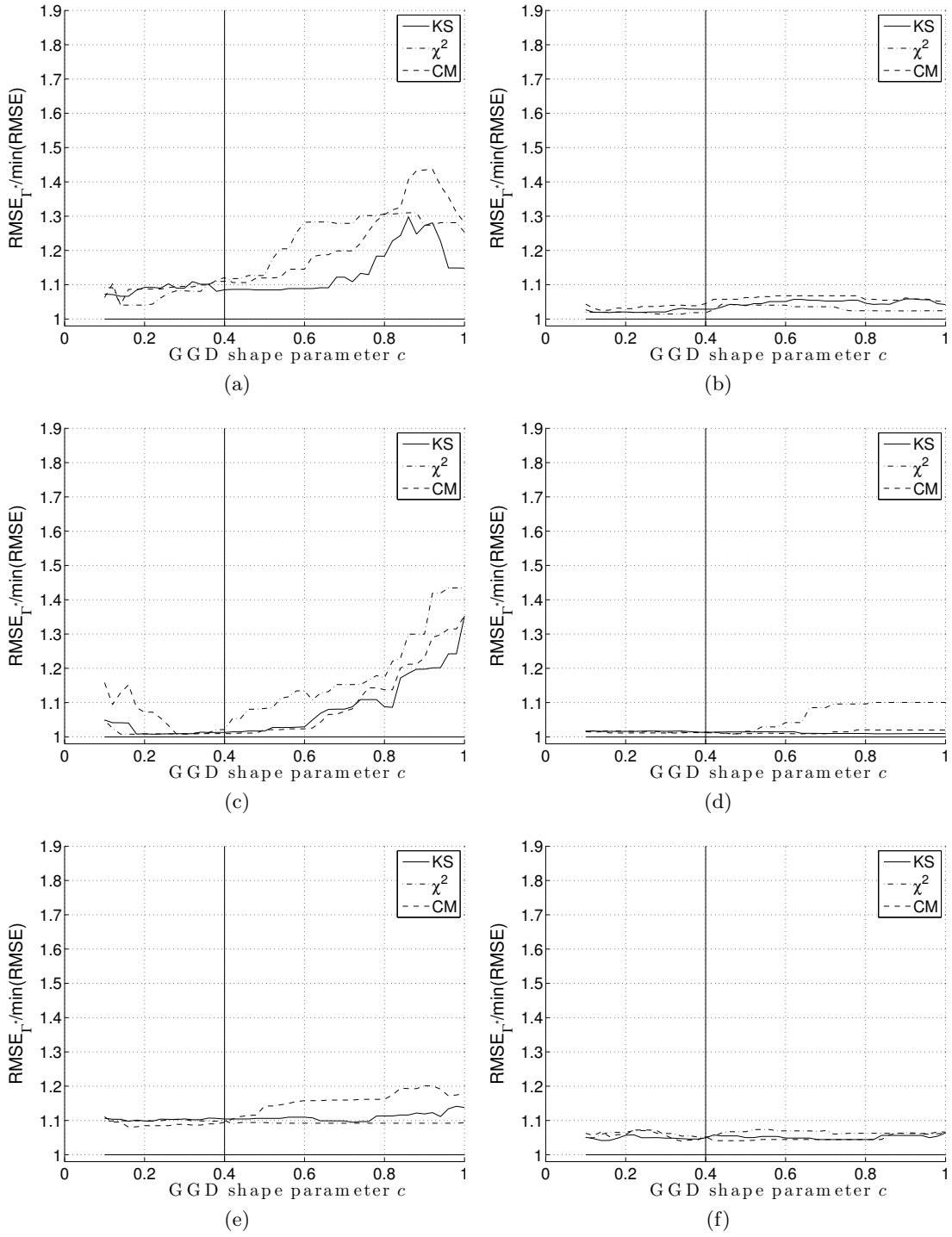


Figure 4.17: Performance of the Γ selection method for 512 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Score functions are averaged over 10 noise realizations and multiplied by the SSE_Γ. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

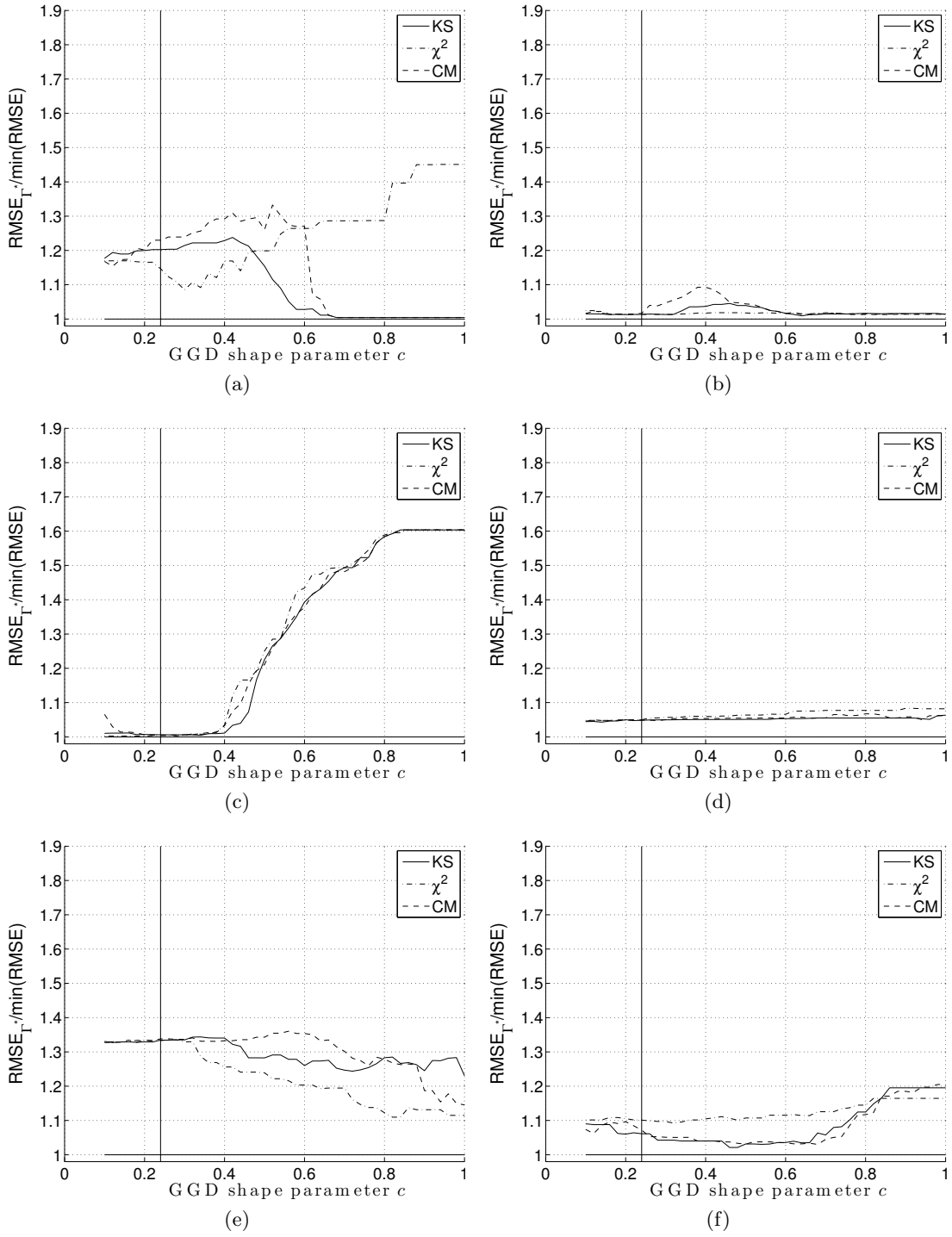


Figure 4.18: Performance of the Γ selection method for 4096 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Score functions are averaged over 10 noise realizations and multiplied by the SSE_Γ. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

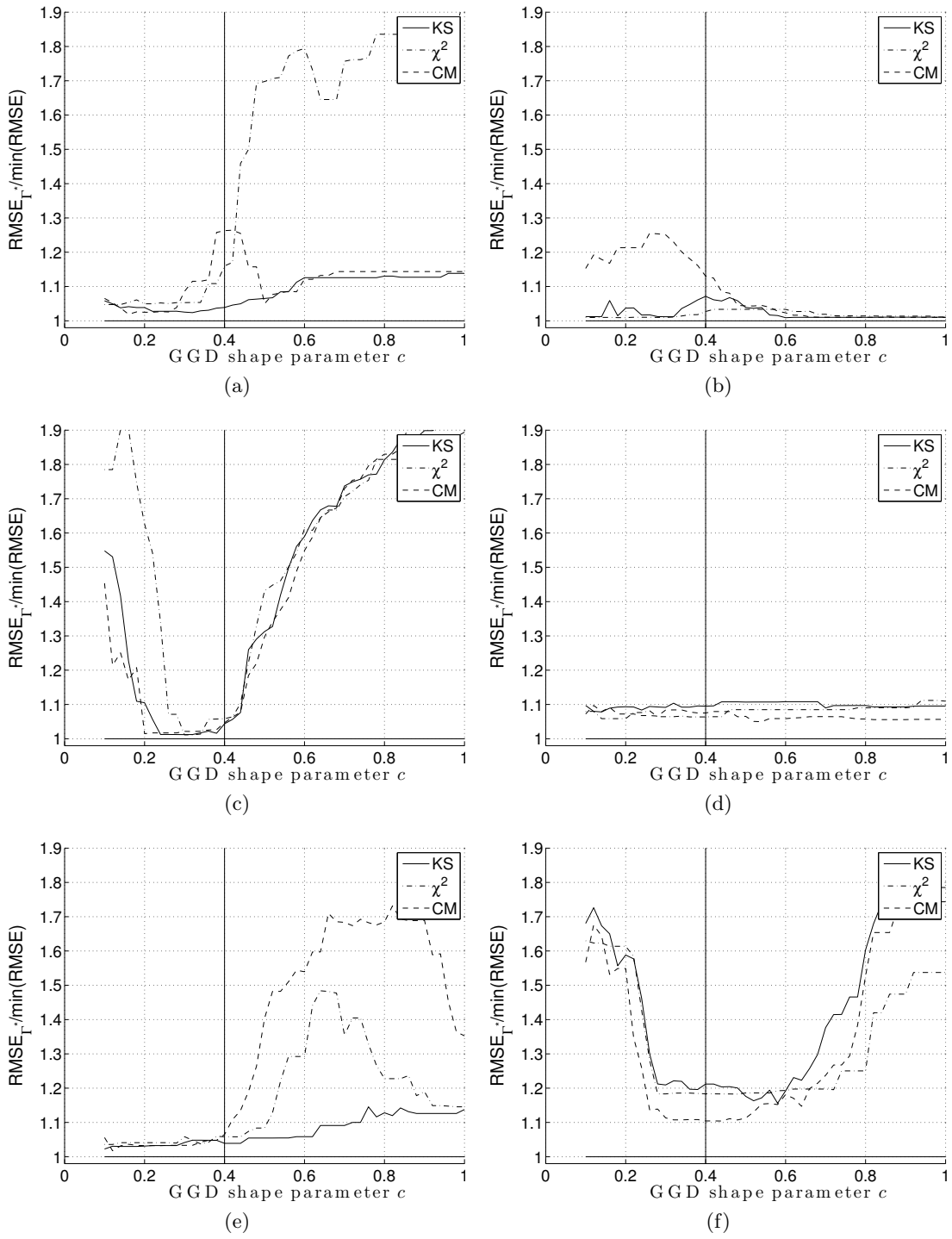


Figure 4.19: Performance of the Γ selection method for 512 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Score functions are averaged over 10 noise realizations and multiplied by the SSE_Γ. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

of the GGD shape parameter c value. In case the 10% ratio was confined to only a few parameter c values, there would be a relatively high probability of significant performance deterioration, after change in any of the influencing factors, described earlier in this chapter, occurs. Thus, we could consider the method to be less stable and provide less reliable results than in the former case.

After analyzing the score functions, we can conclude that the KS score functions are less dependent on the GGD shape parameter c value, than the other two score functions. They also, often, have their minimum extended over a wider range of shape parameter values. In cases in which the $RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o}$ ratio notably rises with the change in c, the rise is often decisively less steep and significant than for the other two tests.

The actual performance of the selection method for the three goodness-of-fit tests can be compared by analyzing Fig. 4.20. For each of the tests, plots of the relative performance ($RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o}$) for all the test cases are shown, along with the histogram of ratios. Histograms clearly indicate that the KS goodness-of-fit test is preferred over the χ^2 and the CM tests. It shows that most of the ratios are concentrated below the 10% mark, and the smallest number of test cases showed unacceptable increase of the RMSE.

Considering that the Kolmogorov-Smirnov test showed advantages over the other two tests, both in terms of the actual performance and the results stability, we can draw the conclusion that it is the preferred goodness-of-fit test for the proposed parameter Γ value selection method.

4.5 Conclusion

It was shown that the ICI-EPL adaptation algorithm can significantly improve signal denoising performance, when compared to conventional wavelet transforms. Still, the algorithm efficiency is heavily dependent on the proper choice of the Γ parameter value. As it is not possible to calculate the value analytically, we proposed two empirical statistical methods for the value selection, from a predefined set of possible values. The methods are based on the analysis of statistical distribution of wavelet coefficients. We showed that if a set of wavelet basis is given, there is a high probability that the basis which is the most suitable for a signal being denoised, will result in a distribution of wavelet coefficients at the last decomposition level which is the narrowest and has the highest peak.

The first selection method is only concerned about the distribution peakedness. The Γ value selection is based on expectation that the distribution of coefficients will be more leptokurtic in case the chosen wavelet basis fits well to a local signal properties, than in the case of an improperly chosen wavelet basis. Although in many test scenarios the expectation was fulfilled and successfully used for the Γ value selection, there were still too many unacceptable results, so we find the method lacking in reliability and efficiency.

The other proposed method is based on modeling the wavelet coefficients distribution at the last decomposition level, using a generalized Gaussian distribution (GGD). The

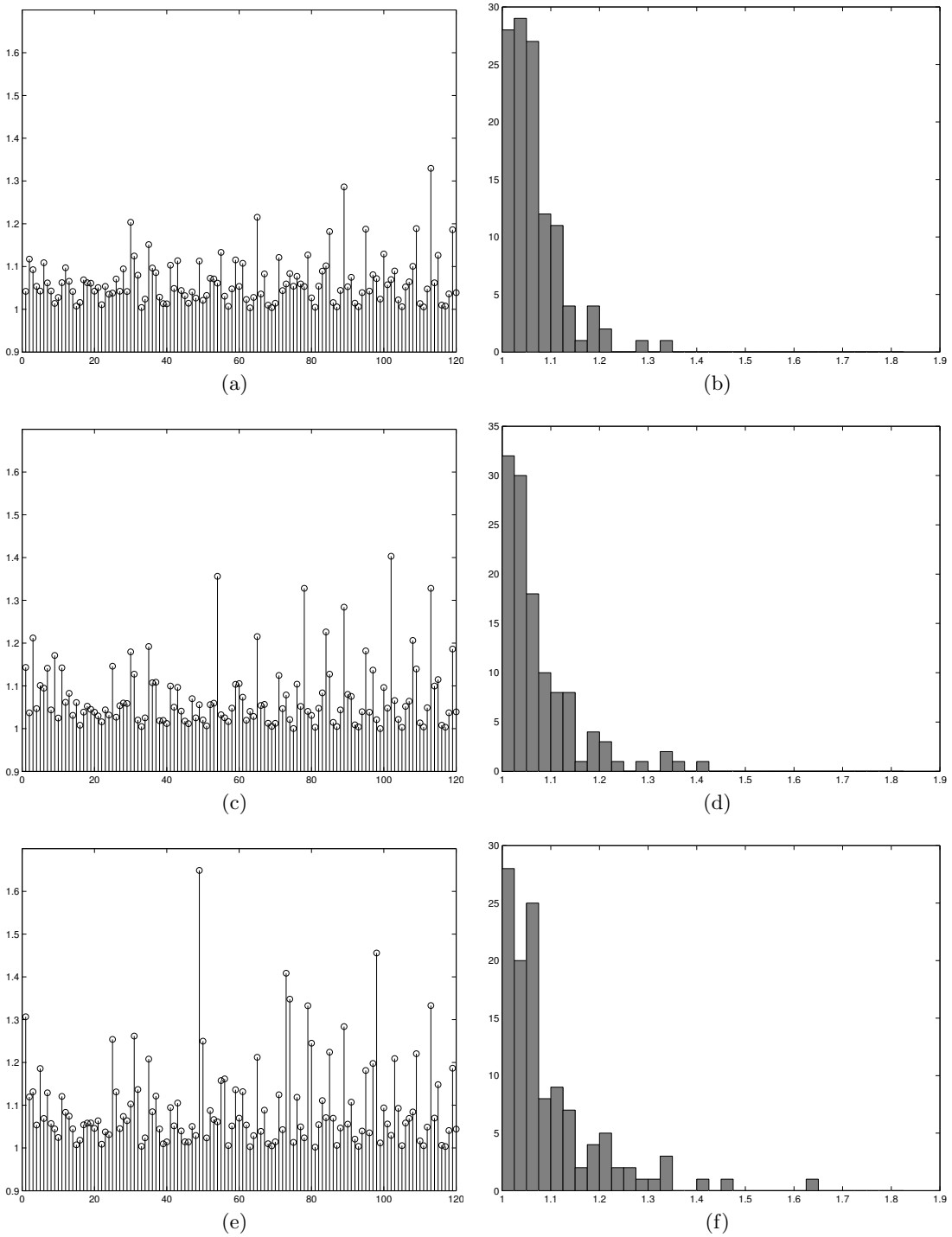


Figure 4.20: $RMSE_{\Gamma^*_{GGD}}/RMSE_{\Gamma_o}$ ratios for each of the test cases, if KS (a), χ^2 (c) or CM (e) goodness-of-fit tests were used. Histogram of ratios for the KS (b), χ^2 (d) and CM (f) tests

strategy for the Γ parameter value selection is to decompose the signal using all available Γ values, and record all the distributions. Each of the distributions is then tested to see how well it fits the GGD, the wavelet coefficients would come close to taking, in the hypothetical case of a perfect fit of the wavelet basis to a signal. Distribution which best fits the GGD is considered to come from the most preferable transform, so the respective Γ parameter value is selected as the most appropriate one.

Three well known methods were used for the goodness-of-fit testing – the Kolmogorov-Smirnov test, the Pearson chi-squared test and the Cramer-von-Mises criterion. In most cases the three tests yielded similar results. When there was greater difference between them, none of the tests proved to yield the best results in each case, but the Kolmogorov-Smirnov test, generally, provided more stable and reliable results than the other two and is, therefore, considered to be the preferred test for the proposed method.

The method was extensively tested on a range of different signal classes, lengths (sampling frequencies) and noise levels. Its efficiency mostly depends on the class of signals being decomposed, while signal length and noise level do not have major influence. The best results are achieved for the *Doppler* signal, for which the average root mean square error (RMSE) degradation is $< 5\%$, when compared to the optimal Γ parameter value case. Typically, it is in the $[2\%, 3\%]$ range. In case of the *Blocks* or *Piece-Regular* signals the average performance degradation is about 10%, but in a few individual cases it can even reach up to $> 30\%$.

We could conclude that, in general, the proposed method performs reasonably well and is suitable for the automated selection of the Γ parameter value, as a part of the ICI-EPL based denoising algorithm. Still, further improvements are possible in this area which would further improve the method efficiency and reliability.

Chapter 5

Application of ICI–EPL to Fluoroscopic Image Sequences Denoising

Conventional surgery procedures assume risks, such as bleeding, cardiovascular risks, or risk of infections. Large incisions that have to be cut cause long post-operative recovery time and increased pain during the gradual healing process. Scars in such procedures are unavoidable and, in some cases, patients decide to undergo cosmetic surgeries to remove them. Because of the negative effects, the conventional surgery procedures are avoided when possible and interventional procedures are preferred and are becoming more frequent. The procedures can most often be considered to be minimally invasive, as only small incisions are cut in the patients skin through which the medical instrument is inserted and further guided through the skin, body cavity or anatomical opening. Depending on the type of the interventional procedure, different imaging techniques may be used to help the surgeon perform the procedure. X-ray fluoroscopy is one such technique, which is used to obtain the X-ray images at high frame rates. It can be used to analyze patients internal structures or to track the position of medical instruments. Some common treatments where fluoroscopy is used include placement of peripherally inserted central catheter, balloon angioplasty, cardiac ablation or implantation of cardiac rhythm management devices, such as pacemakers. The biggest concern of the technique is the cumulative radiation dose delivered to the patient. A single frame dose is very low, however, procedure times of up to 75 minutes have been recorded. Such prolonged exposures may eventually lead to a severe skin injuries (Berlin, 2001; Frazier, Richardson, Fabre, and Callen, 2007; Koenig, Mettler, and Wagner, 2001). Lowering the single frame dose is of the utmost importance but lower dose also implies heavier noise image in the image. Too high noise level could render the X-ray image sequence unusable during the procedure, so a trade-off has to be made between the exposure and the acceptable image noise level. Efficient image processing algorithms can help in removing the noise and

improving image quality, thus, allowing for further dose reduction.

Many methods have been developed to address the problem of noise in a sequence of X-ray fluoroscopy images. Aufrichtig and Wilson (Aufrichtig and Wilson, 1995) take the object-detection approach and assume that most of the motion is constrained to a long thin objects, such as catheters or guide wires. The method uses small oriented line segments as template filter kernels to discover such objects. Schoonenberg et al. (Schoonenberg, Schrijver, Duan, Kemkers, and Laine, 2005) focus on distinguishing the moving objects (human organs) from the static background as well as on the catheter detection. They propose adaptive filtering process in which motion compensated temporal filtering is applied to moving parts of the image while spatial filtering is applied to static image segments. Special care is taken on occasional catheter jumps between the two subsequent frames. Manjeshwar and Dhawale (Manjeshwar and Dhawale, 2005) achieved good results in image quality improvement by using a method based on the motion discriminating temporal filter (MD-TF), which reduces motion-blur typical for some commonly used temporal recursion filters. Their findings were assessed by both the realistic image synthesis model and the visual perception experiments. Bismuth and Vaillant (Bismuth and Vaillant, 2008) processed fluoroscopic image sequences in interventional cardiology. Their algorithm is based on linear feature detection in which they try to distinguish medical instrument from the background. The final image is composed of denoised background and enhanced linear features.

In this chapter, we propose another method for denoising the fluoroscopic image sequences. It is not tailored for a specific treatment but should be applicable to any image sequence presenting a nearly static background, with the motion confined to a human tissue, organ or medical instrument. Instead of the feature/object detection route, redundant information stored in the temporal dimension is heavily exploited. Time window of images is considered and two estimates of denoised images calculated. One is obtained by transforming a set of images to a 1-D signal by 3D scanning of image pixels, and denoising the resulting signal using the edge preserving wavelet transform ICI-EPL, presented in Chapter 3 and (Tomic, Sersic, and Vrankic, 2008). The other estimate is obtained by applying a basic intersection of confidence intervals (ICI) algorithm (Katkovnik, 1999) to the temporal dimension of the time window image set, in order to further improve performance near object edges. Two estimates are fused using statistical properties of the ICI estimators, yielding high quality denoised images, with preserved edges between objects.

5.1 Exploring the Temporal Dimension I – Wavelet Transform

Over time, many image denoising methods were developed, largely focused on natural images. Their efficiency, most often, rely on the well known fact that the human visual

system (HVS) is much more susceptible to lower frequencies than to higher ones. Image compression algorithms utilize the fact by removing the higher frequencies from the image, allowing for image description with fewer bits. Image denoising algorithms also rely on the same HVS property. Noise in natural images represent the high frequency component, so removing the higher frequencies will also cause the noise to be attenuated, yielding images of much improved perceived quality. In the case of medical images, it is a common situation that the higher frequencies carry valuable information for medical personnel. Quite often, large low frequency regions have little meaning, while higher frequency regions, such as edges between objects, human tissue or blood vessels are more important. It is the opposite assumption to the one taken for natural images. It was shown that the ICI-EPL performs well around edges in a signal, and it is to be expected that it will also perform well on medical images and tend to efficiently remove the noise, while still reconstructing the intrinsic image higher frequencies. The requirement is essential for application to a fluoroscopic image sequences.

As presented, ICI-EPL is proposed as a 1-D transform. Since it was a wavelet transform, it is straightforward to generalize the method to multiple dimensions by means of separable filter banks. The 3-D generalization can be achieved by addition of the filtering in the temporal direction. Because of the specifics of the given application, neither 2-D nor 3-D generalization could be considered to be an optimal solution. In the 2-D case, the temporal dimension is neglected and each X-ray image is denoised as if it was independent of the images that preceded it. As subsequent images often differ in only a few details, ignoring this redundancy cannot be considered to be the best approach. On the other hand, the 3-D generalization of the ICI-EPL would quite efficiently use all the spatial and temporal data available. However, there are two problems associated with the approach, and both of them stem from the necessary amount of data in the temporal dimension. Properties of the wavelet transform could not be utilized to a desirable extent if there were less than 3 levels of wavelet decomposition, which again presumes approximately 256 subsequent images. The two problems introduced with such large data sets are excessive delay and long computation times. Even if the 3-D approach performed well in *off-line* denoising, in case the real-time denoising requirement could not be satisfied, it would be pointless.

Described multidimensional generalizations using separable filter banks assume that the original 1-D transform has to be modified in order to efficiently process multidimensional signals. For this particular application, the optimal solution is searched for by using the opposite approach. Instead of generalizing the transform to fit the multidimensional signal, the image sequences are modified and "fitted" to the 1-D transform.

5.1.1 3-D Sample Scans

To process the set of images from a sequence with a 1-D transform, the 3-D data has to be converted to 1-D data. It is done by means of the 3-D sample scanning. Traversing

the data in the horizontal, vertical and zig-zag 3-D paths, three 1-D signals are produced. The horizontal and the vertical 3-D scans are depicted in Fig. 5.1. In the horizontal case, pixels are scanned back and forth along temporal dimension, with horizontal offset at the ends of the image set. After the first row is fully scanned, the scan continues at the next row. In the vertical case, the offset is vertical and after the first column is fully scanned, the scan continues at the next column.

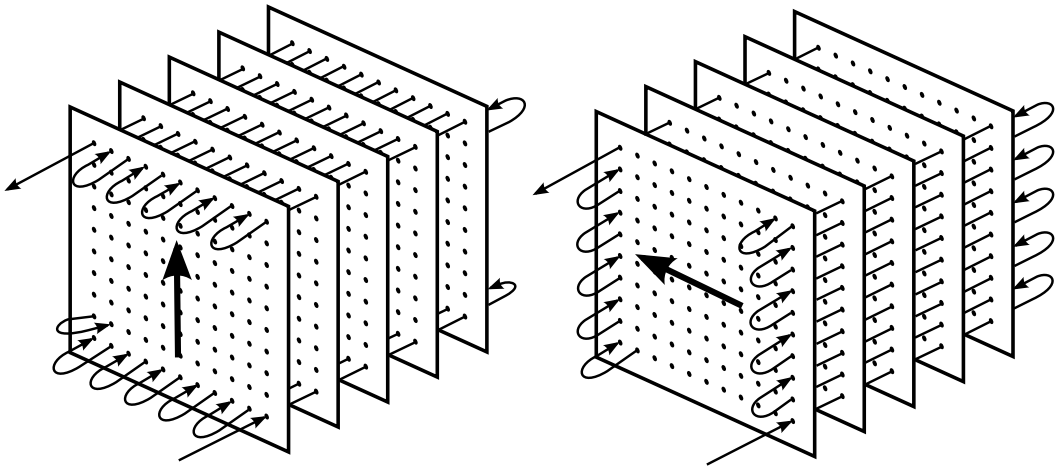


Figure 5.1: Horizontal (a) and vertical (b) 3-D scan of a set of subsequent fluoroscopic X-ray images

In the zig-zag case the traversing progresses in a zig-zag fashion, in a diagonal direction. The 2-D zig-zag scan is depicted in Fig. 5.2 and it's 3-D implementation is constructed following the same principles as in the horizontal or vertical case.

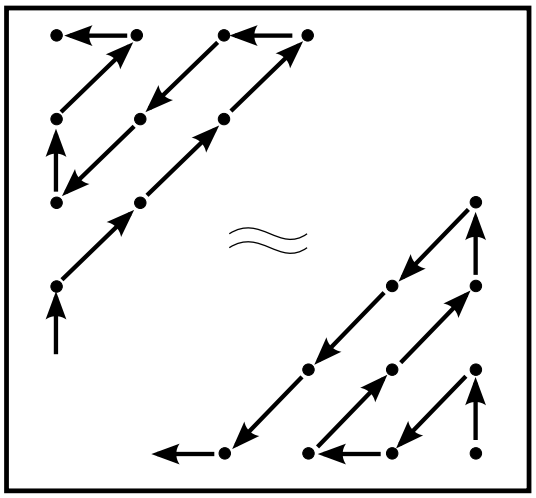


Figure 5.2: 2-D zig-zag pixel scan

The conversion of image sets to a 1-D signal allows for efficient utilization of the temporal dimension using wavelet transforms, while still requiring only a small number of

images, unlike the regular 3-D transform. It is to be expected that pixels in subsequent images, which share the same coordinates, will in most cases have close values, with difference often originating from noise. The same is often also true for surrounding pixels of any pixel in a single image. The idea of the 3-D scans is to bring these pixels together. The resulting 1-D signal should consist of many rather smooth regions with only minor pixel value fluctuations caused by a noise. Edges between the regions are to be emphasized and clearly identifiable. An example of such 1-D signal is shown in Fig. 5.3.

The set-up provides convenient input for the ICI-EPL, as smooth regions can be denoised efficiently, while edges should be pronounced enough to allow for good transform adaptivity. The assertion is best explained in the extreme case of an image containing a single pixel wide line. Should the image be denoised using the 2-D separable generalization of wavelet transforms, there is a high probability of the line completely disappearing from the denoised image. Given the 3-D scan of, for instance, 6 subsequent images, the same single pixel line would actually be 6 pixels wide and have much better chance of survival in the denoised image.

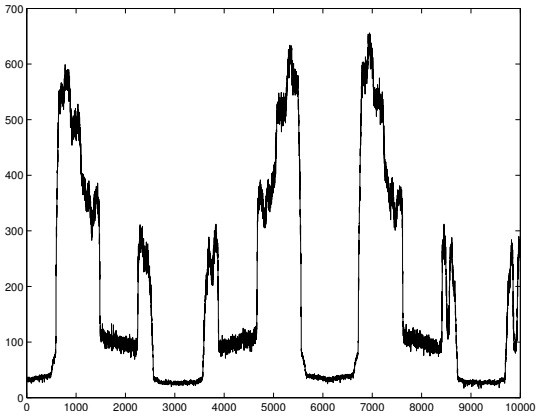


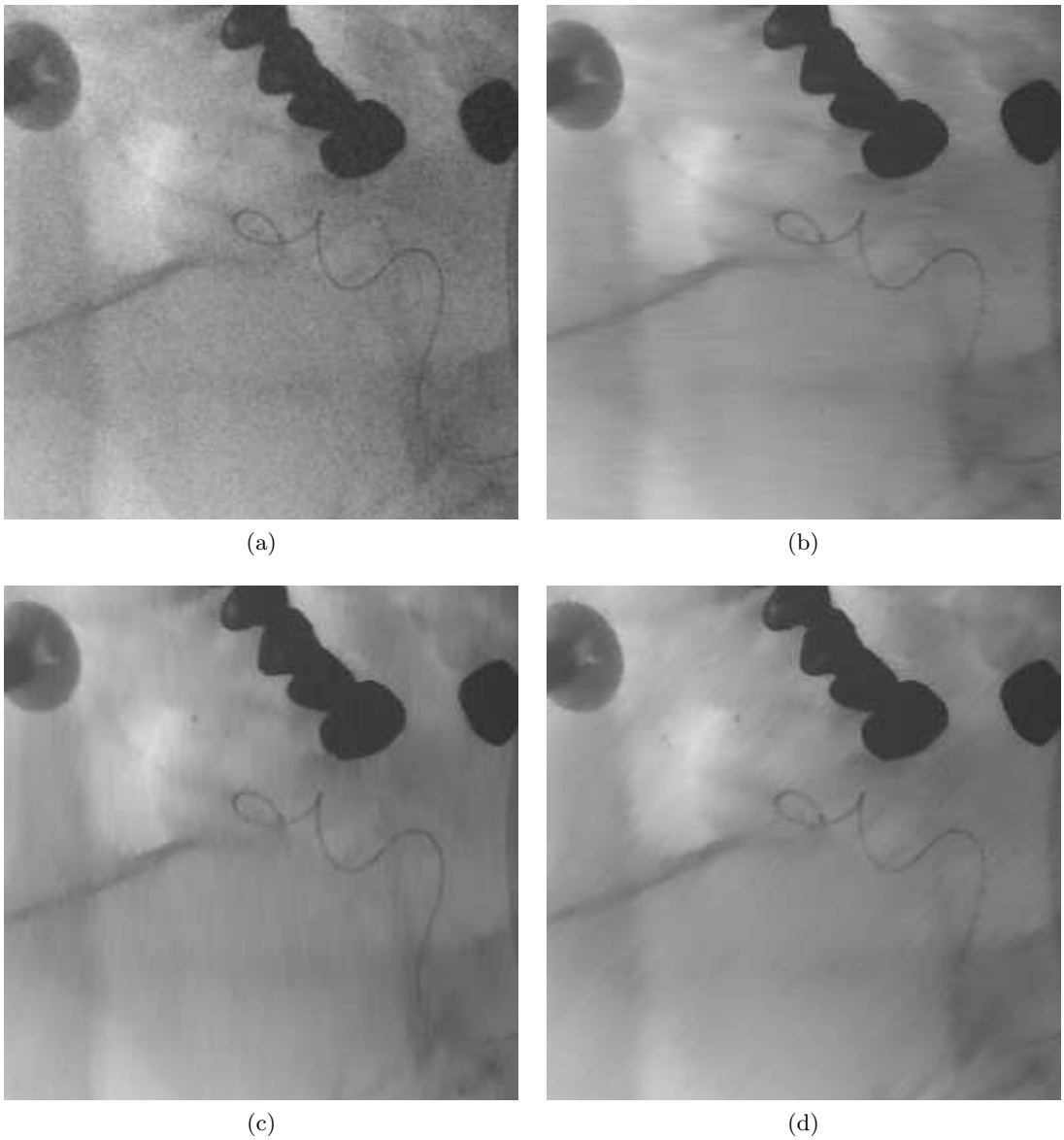
Figure 5.3: Part of the sample 1-D signal, resulting from the 3-D horizontal scan of 6 subsequent images

5.1.2 Combining the Scans and Generating Denoised Image Estimate

After the conversion to a 1-D signal, each of the three signals is independently denoised using the unmodified ICI-EPL 1-D transform and converted back to the 3-D by inverse scanning. Sample noisy image and three denoised images are shown in Fig. 5.4. Only two subsequent images were used for denoising. Resulting images suffer from clearly visible horizontal, vertical and zig-zag artifacts, in respective images. The artifacts cannot be avoided because images were not processed by a real 2-D or 3-D transform. To better understand the effect of processing the image set using the 1-D transform, it is best to analyze the wavelets used in signal reconstruction, and their dispersion during the inverse scanning. For the purpose, fixed wavelet *Bior2.2* is used.

In 1-D, the *Bior2.2* wavelet looks as depicted in Fig. 5.5. After inverse scanning the

1-D signal, wavelet disperses over all images in a set. The dispersion for several set sizes is shown in Fig. 5.6. In case of a single image in a set, there is no dispersion and entire wavelet is contained in that single image. Depending on the type of scan, the wavelet stretches in either horizontal, vertical or diagonal direction. As the number of images in a set rises, the wavelet disperses over all of the images and the prevailing direction of the stretch becomes less pronounced. In effect, the artifacts in the denoised image slowly diminish, although do not entirely disappear. To minimize artifact visibility, regardless of the wavelet transform being used, the three images are averaged to form the single denoised image.



(a) (b) (c) (d)

Figure 5.4: Noisy image of catheter insertion (a) and denoised images resulting from 1-D signals obtained by horizontal (b), vertical (c) and zig-zag (d) 3-D scans of two subsequent images

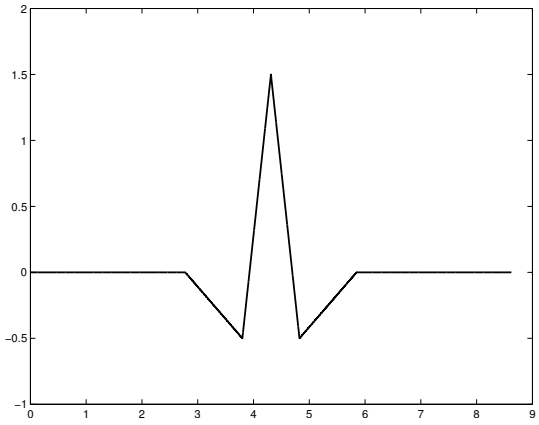


Figure 5.5: Bior2.2 synthesis wavelet function

The proposed approach of fitting the signal to the transform has two advantages. One is that it allows for very efficient utilization of temporal dimension while still heavily relying on each image spatial data (as can be seen in Fig. 5.6). This way, the advantages of the full 3-D approach are inherited but with much smaller image sets. The smaller image set size leads to the other advantage - the transform can be calculated much faster than it would be possible in a 3-D case, making it more suitable for real-time applications.

5.2 Exploring the Temporal Dimension II – the Basic ICI

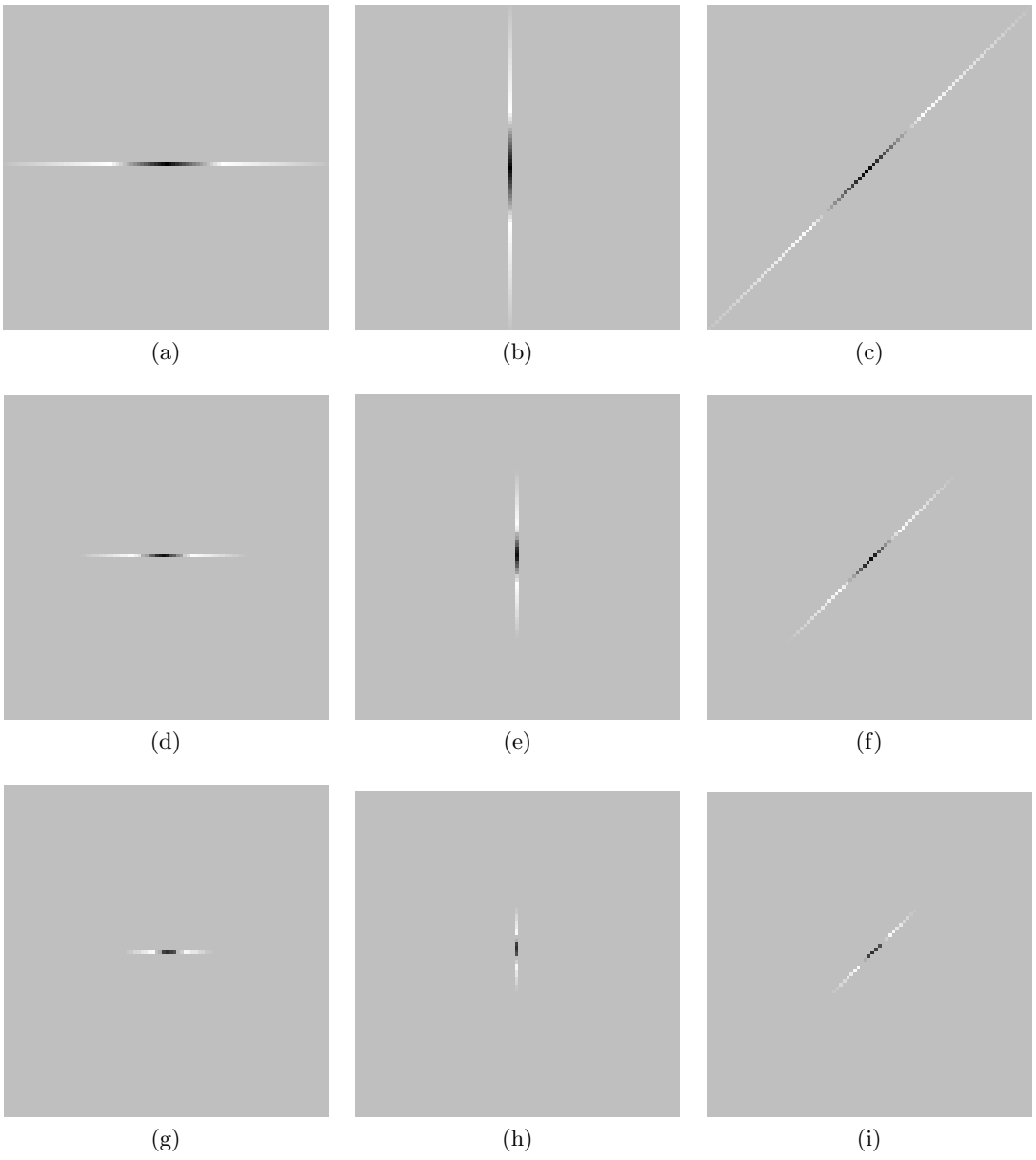
ICI-EPL provides very good results in denoising the fluoroscopic images, however, given the significance of edge retention, one more estimate of denoised images is introduced. It is based on the simple implementation of the ICI rule Katkovnik (1999), presented in more details in Chapter 3.3. The basic ICI algorithm is applied along the temporal dimension of a set of images, for each pixel independently. As piece-wise constant model of the signal is assumed, the averaging filters may be used as a set of estimators of growing supports:

$$\hat{y}_{h_j}(\boldsymbol{x}, i) = \frac{1}{j} \sum_{k=0}^{j-1} z(\boldsymbol{x}, i-k), \quad j = 1 \dots N, \tag{5.1}$$

where $\boldsymbol{x} = (m, n)$ are spatial coordinates of a given pixel, i is ordinal number of image in a sequence and N is total number of images in a set. Although the scheme alone cannot provide high-quality denoised image sets, it is efficient at preventing the blurring of edges between objects and it is used to enhance the quality of denoised image estimates obtained by the ICI-EPL.

5.3 Fusion of the Estimates

There are multiple ways of fusing the two estimates. Estimators used to produce the estimates (adaptive wavelet filter banks and the ICI algorithm) differ in their statistical

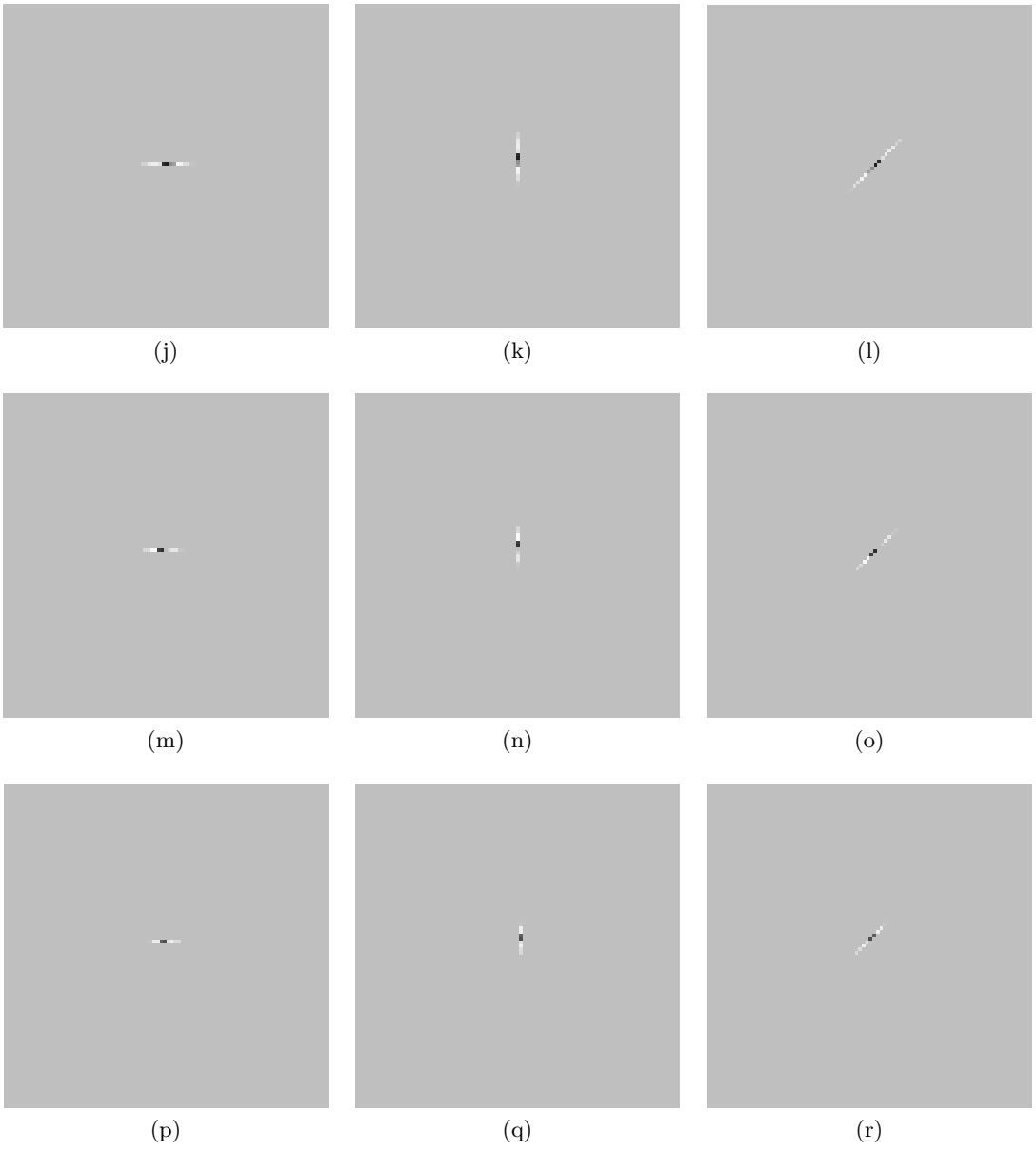


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 5.6: Wavelet dispersion after inverse scanning the 1-D signal. Each row shows wavelets after inverse horizontal, vertical and zig-zag scan. Number of images in a set is: 1 (a)-(c), 2 (d)-(f), 4 (g)-(i)



(j) (k) (l)

(m) (n) (o)

(p) (q) (r)

Figure 5.6: Wavelet dispersion after inverse scanning the 1-D signal. Each row shows wavelets after inverse horizontal, vertical and zig-zag scan. Number of images in a set is: 6 (j)-(l), 8 (m)-(o), 10 (p)-(r)

properties. Because of that, ordinary mean would not produce the best results. In general, optimal results are obtained by using the weighted mean, where weights are based on the reciprocal of the estimate variance:

$$\hat{y}(\boldsymbol{x}) = \frac{1}{\sum_l 1/\sigma_l^2} \sum_l \frac{\hat{y}_l(\boldsymbol{x})}{\sigma_l^2}, \tag{5.2}$$

where index l corresponds to the l-th estimate. Given that the estimators are FIR filters, the variance of each pixel is given by:

$$\sigma_{\hat{y}_{\boldsymbol{x}}}^2 = \sigma_\epsilon^2 \cdot \sum_k g_{h_{\boldsymbol{x}}}^2(k), \tag{5.3}$$

where σ_ϵ is noise variance and $g_{h_{\boldsymbol{x}}}$ is impulse response of the estimator used for the pixel at position $\boldsymbol{x} = (x_1, x_2)$. From (5.2) and (5.3) follows:

$$\begin{aligned} \hat{y}(\boldsymbol{x}) &= \frac{1}{\sum_l w_l} \sum_l w_l \hat{y}_l(\boldsymbol{x}), \\ w_l &= \frac{1}{\sum_k g_{lh_{\boldsymbol{x}}}^2(k)} \end{aligned} \tag{5.4}$$

For averaging filters used in the ICI estimate, it is straightforward to use (5.4), with weights being $w_I \in 1, 2, 3, \ldots, N$. Since ICI-EPL is a highly adaptive wavelet transform, estimating exact weights for each pixel, based on (5.4), would be a difficult and costly operation. Therefore, setting the weight for the ICI-EPL estimate to a fixed value is proposed:

$$w_{IE} = const. \tag{5.5}$$

Choice of the w_{IE} decides the impact of the ICI–EPL estimate to a final denoised image. The ICI-EPL estimate is considered to provide the base of denoised image, which is only further enhanced by the ICI estimate. To promote its significance, we let its weights to be equal to the weight of a reliable ICI estimator. Range of reasonable values can only be determined empirically and we find it reasonable to set the w_{IE} to a weight of ICI estimator with support length of 10 pixels:

$$w_{IE} = 10.$$

Although a sub-optimal solution, it still allows for efficient fusion, i.e. variance of the fused estimate should be lower than the variance of either of the estimates. As well, the final image should give higher PSNR value then either of the estimates. Another benefit of the fixed-weight approach is the ability to easily adjust the general impact of the ICI-EPL estimate in the final image.

5.4 Results and Discussion

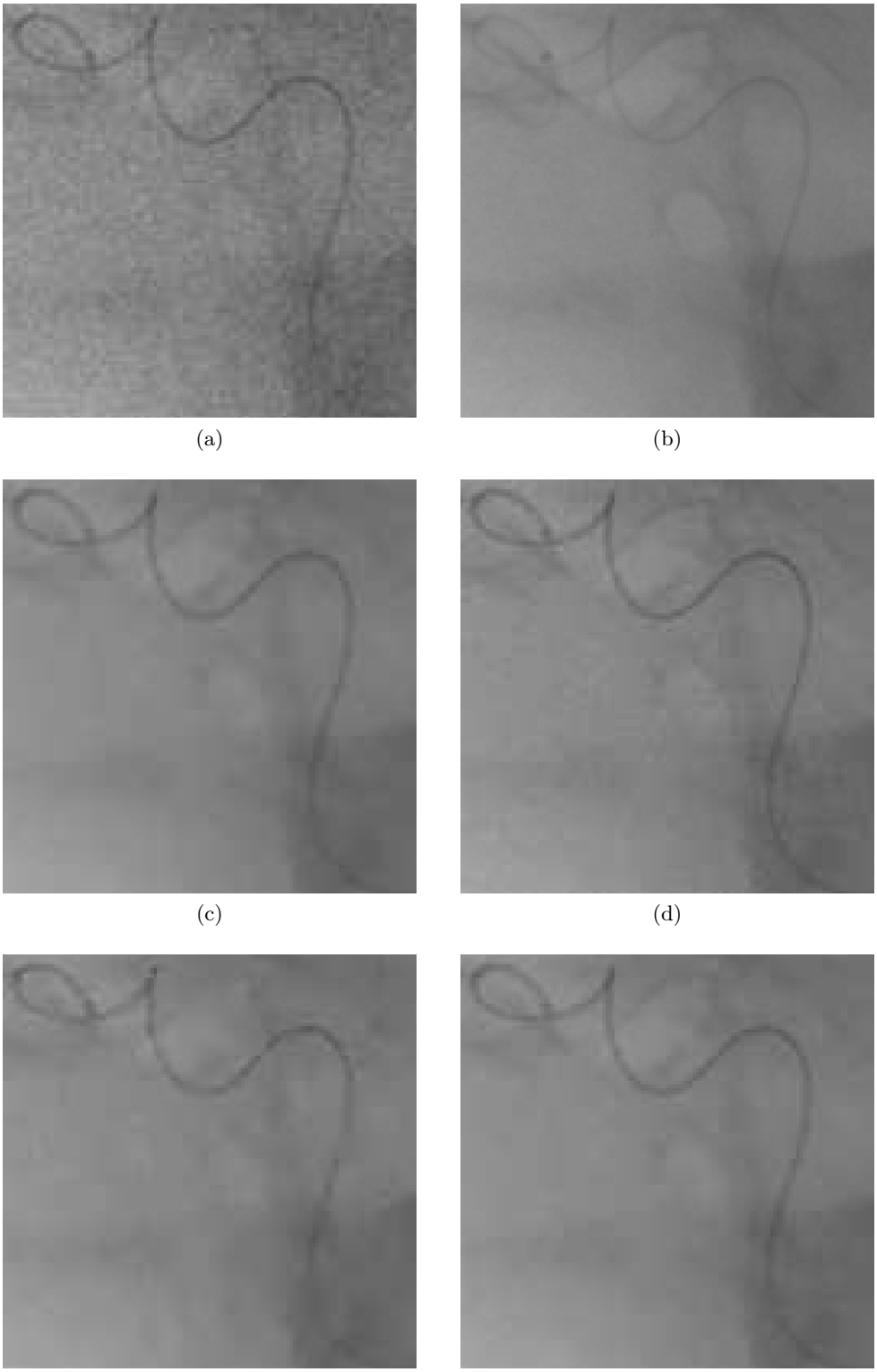
The effectiveness of the proposed method has been tested on a real fluoroscopic image sequence of catheter insertion. Resolution of images is 512x512 pixels and they were taken at 10 bits per pixel precision. For the wavelet transform estimate, four different transforms were used: *Haar*, *Bior2.2*, *Bior4.4* and ICI-EPL. The method is applied on image sets of 4, 6, 8, 10 and 12 consecutive images. In case of the ICI-EPL, denoising was performed, and performance examined, for a full range of reasonable Γ parameter values: $[0.5, 5]$. The Γ value for the ICI step is always set to a value in the lower half of the range: $\Gamma = 1.7$.

As a general denoising performance measure, PSNR values are given in Table 5.1. As a clean signal approximation, necessary for the PSNR calculation, we averaged 250 consecutive sequence images. It can be seen that the ICI-EPL estimate yields better estimate than the basic ICI. The total gains of $10dB$ are achieved, with the fusion of the two estimates adding up to $0.5dB$ to the ICI-EPL estimate PSNR. Although seemingly small difference, the fusion step is still very important and, as will be shown in the next example, it is probably not reflected by the small $0.5dB$ improvement in the PSNR. The significance of the improvement achieved by the fusion is in the the edge enhancement, as edges are the most important part of the image to the medical personnel. This is better observed in Fig. 5.7, in which results of applying the method to a set of 8 consecutive sequence images are shown.

It can be seen that the ICI-EPL estimate is of high quality. The noise is very efficiently removed and edges between the catheter and the background are mostly well preserved, however, still slightly blurred in certain areas. Edges in the ICI estimate are better pronounced than in the ICI-EPL estimate, but significant level of noise remained existent in the denoised image. Finally, the fused estimate combines advantages of both of the compounding estimates and provides denoised images, in which catheter and its tip are clearly and easily distinguished from the background. Showed for comparison is also the 2D separable implementation of the ICI-EPL applied to a single image. As expected, it cannot compete with the proposed spatio-temporal method.

To explore potential benefits of the proposed method, we introduce additional Gaussian noise to the examined image sequence. The results of denoising are shown in Fig. 5.8. In the previous example, even the noisy image could have been considered to be quite usable, but after artificially raising the noise level, noisy image becomes completely useless as it is impossible to discern the catheter from the noise. Sequence was denoised using image set size of 8 subsequent images. Naturally, denoised images are of much lower quality than in Fig. 5.7. There is more blurriness at the catheter tip as well as the wire but it can also be seen that both the catheter wire and its tip can still be clearly distinguished from the background.

PSNR values for denoised images can be found in Tab. 5.2. Denoising algorithm brought improvements of $16dB$. It is interesting to note that the fused image PSNR can



(a) (b) (c) (d) (e) (f)

Figure 5.7: Denoising results for catheter insertion sequence. Zoomed in: noisy image (a), approximation of clean image (b), ICI-EPL estimate of denoised image (c), basic ICI estimate of denoised image (d), image denoised by separable 2-D implementation of the ICI-EPL (e) and fused estimate (f)

Table 5.1: PSNR values of denoised images for image sets of 2, 4, 6, 8, 10 and 12 consecutive images

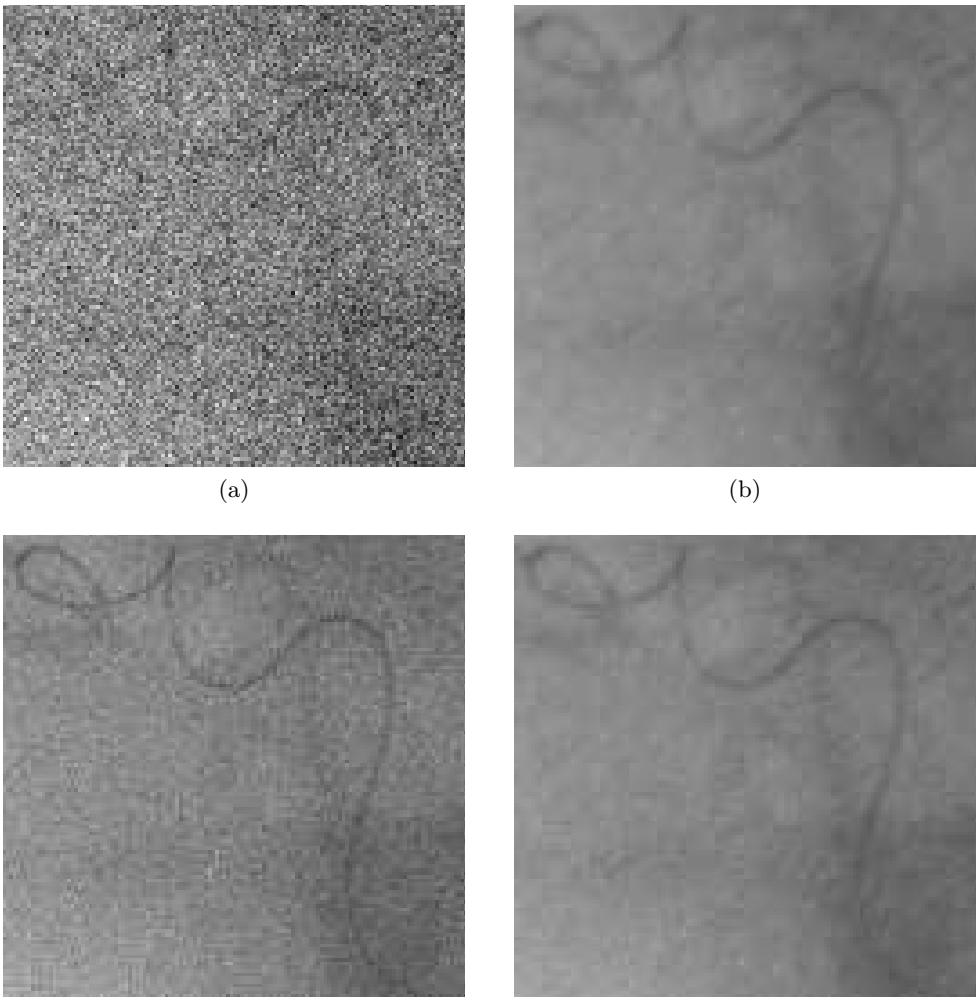
	2	4	6	8	10	12
Noisy	40.01	40.01	40.01	40.01	40.01	40.01
ICI	42.62	45.22	46.65	47.55	48.08	48.55
Haar	46.90	48.23	48.94	49.37	49.69	49.96
Fused Haar	46.93	48.40	49.20	49.73	**50.10**	**50.43**
Bior 2.2	**47.04**	48.12	48.97	49.00	49.25	49.43
Fused Bior 2.2	**47.04**	48.29	48.67	49.41	49.73	49.99
Bior 4.4	46.80	48.05	48.7	49.64	49.41	49.62
Fused Bior 4.4	47.00	48.4	49.13	49.13	49.96	50.23
ICI-EPL	46.74	48.21	48.96	49.39	49.58	49.82
Fused ICI-EPL	46.82	**48.47**	**49.25**	**49.75**	49.99	50.28

get up to $1.5dB$ lower than the PSNR of the wavelet transform based estimates (ICI-EPL and conventional wavelets). It is a typical example of the insufficiencies of the PSNR as a generally applicable objective measure of denoised image quality. If we compare denoised images before and after the fusion (Fig. 5.8 (b) and (d)) it is evident that the catheter and its tip are less blurry and better distinguished from the background in the fused image than in the ICI-EPL estimate. As medical personnel focus on the catheter, we find the fused image to be more usable in clinical application. Such specifics of the given application cannot be addressed by the PSNR measure.

Table 5.2: PSNR values of denoised images for image sequence with raised noise level

PSNR Noisy	25.95			
PSNR ICI	34.78			
	ICI-EPL	Haar	Bior2.2	Bior4.4
PSNR Wavelet	42.25	42.95	42.34	42.25
PSNR Fused	**40.92**	41.34	**40.92**	40.93

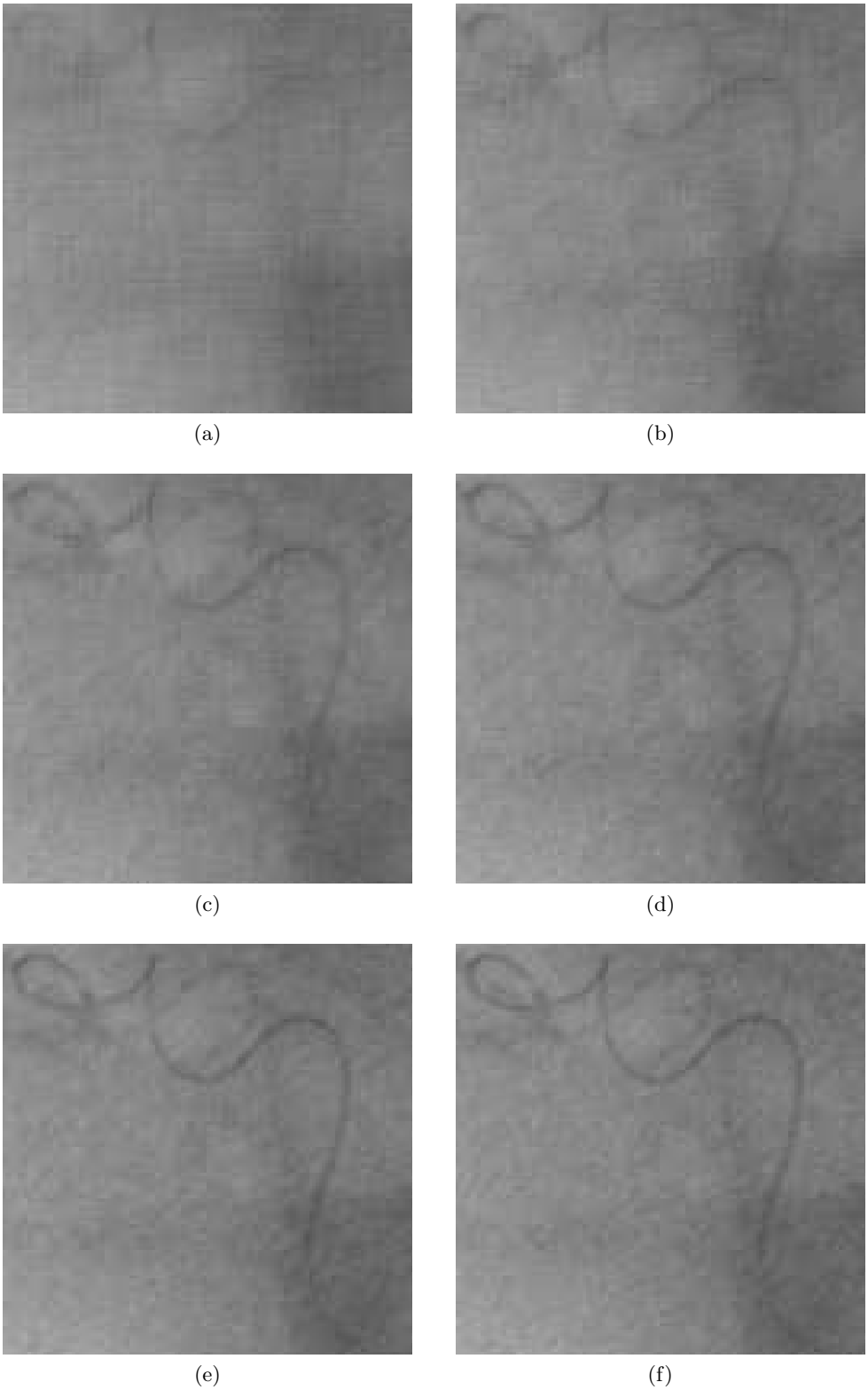
Fig. 5.9 shows the improvement of fused denoised image quality as the number of images in a set rises. Generally, larger image sets result in better quality of denoised images, but we must also, again, consider some specifics of the given application. Enlarging the image set adds to processing time, which is a limiting factor since this is a real-time application. Also, as the number of images in a set rise, so does the effect of averaging. Quality of the background indeed improves, but quality of moving objects (catheter) in an image does not or it could even deteriorate. We find that the reasonable number of images in a set is 6-10. Larger image sets show some improvement in terms of the PSNR, but the real clinical benefit or detriment introduced by such sets would be difficult to evaluate without the real-world clinical or visual perception experiments.



(a) (b)

(c) (d)

Figure 5.8: Denoising results for catheter insertion sequence with artificially added noise. Zoomed in: noisy image (a), ICI-EPL estimate of denoised image (b), basic ICI estimate of denoised image (c) and fused estimate (d)



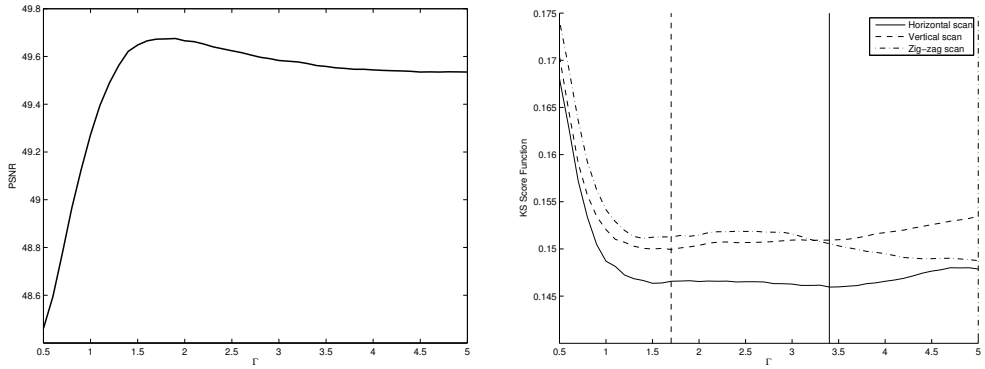
(a) (b) (c) (d) (e) (f)

Figure 5.9: Denoising results for catheter insertion sequence for image sets of : (a) 2, (b) 4, (c) 6, (d) 8, (e) 10 and (f) 12 images

5.4.1 Automated Γ parameter selection

All the results in this chapter were obtained using the optimal value of the Γ parameter for the ICI-EPL estimate, Γ_o. As it is not possible to calculate this exact value, we should examine the dependency of the denoised image PSNR on the chosen value for the Γ parameter, and the efficiency of the algorithm for automated Γ parameter selection, devised in Chapter 4.

Fig. 5.10 (a) shows PSNR of the denoised image plotted against the full range of Γ parameter values. Unlike in the 1-D case, selection of the Γ parameter does not influence the denoising results significantly. Difference between the best and the worst case performance is only slightly more than $1dB$. Also, for any chosen Γ parameter value in the range $\Gamma \in [1, 5]$ the PSNR does not deteriorate more than $0.5dB$. Fig. 5.10 (b) shows score functions for the Γ selection algorithm. Separate score function is plotted for each of the scans, with vertical lines marking the functions minimum, i.e. the chosen Γ value: 1.7, 3.4, 5. In each case, the difference in PSNR when using selected and optimal Γ value for denoising is $< 0.2dB$. If we average the selected values, the final Γ parameter value becomes $\Gamma^* = 3.36$, resulting in a $0.1dB$ lower PSNR than for the optimal Γ value.



(a) (b)

Figure 5.10: Dependency of denoised image PSNR on the chosen Γ parameter value (a) and score functions for automated Γ parameter selection (b)

5.5 Conclusion

X-ray fluoroscopy is common method in interventional procedures. Since the procedure can take more than an hour to complete, the problem of cumulative radiation dose to the patient can become very significant and even lead to severe injuries. The only two solutions to the problem are shortening the procedure time and lowering the fluoroscope single frame dose. While the procedure time may be shortened slightly, as a result of surgeon experience or development of new tools, realistically, only the lowering of the single frame dose can lead to a substantial lowering of the cumulative radiation dose. However, if the single frame dose is lowered too much, the images become unreadable

because of the too low signal-to-noise ratio. Denoising algorithms can be used to process the images and make them usable, even though they were taken with a very low single frame dose.

Denoising of X-ray image sequences using pure 2-D image denoising algorithms cannot be considered as an optimal approach to the problem. Redundancy of certain image parts in the temporal dimension can be efficiently explored to improve denoising algorithm performance, so spatio-temporal denoising algorithms are preferred way of obtaining high quality denoised images. Another specific of this application is treatment of edges in an image. In X-ray fluoroscopic sequences such sharp edges carry valuable information for the medical personnel so it is essential to prevent blurring, or even, blending them with the neighboring objects or background.

The method we propose relies on extensive use of temporal dimension to yield denoised images in which objects edges do not get blurred, but stay clearly identifiable and distinguishable from the background. The idea of the algorithm is to produce two estimates of denoised signal.

One is obtained by transforming the set of images from the image sequence to a 1-D signal, and then perform denoising using the ICI-EPL, as if it was an ordinary 1-D signal. Set size is arbitrary, with larger sets leading to better denoising performance on the expense of prolonged processing times. Too large sets can also have other adverse effects. Results showed that sets larger then 10 images should not be necessary. After denoising, the signal is transformed back to 3-D, resulting in the ICI-EPL estimate of denoised signal. The estimate is, generally, of high quality, with edges mostly well preserved.

The other estimate is based on applying the basic ICI algorithm to the same image set. Algorithm is applied in the temporal dimension and for each pixel independently. The ICI estimate is, generally, of much lower quality, but retains the edges better. The two estimates are fused to produce the final denoised image. The fusion might be viewed as an edge enhancement step for the ICI-EPL estimate.

The method was tested on a real fluoroscopic image sequence of catheter insertion. The results show that the spatio-temporal implementation of the ICI-EPL can be used to efficiently denoise X-ray fluoroscopic image sequences. Fusing the ICI-EPL image with the basic ICI denoised image further enhances the edges between objects and improves image readability. Simulations were also performed on image sequences with artificially raised noise level. Such images, when unprocessed, were completely unreadable as it was impossible to discern any of the objects from the noise. After denoising with the proposed spatio-temporal method the catheter wire and its tip became clearly visible.

The quality of the final denoised images suggests that the proposed algorithm could allow for significant lowering of the fluoroscope single frame dose, leading to a reduction in the cumulative radiation dose to the patient. However, it would be very difficult to explore the full potential of the proposed method without the real-world tests. Visual perception experiments should be carried, in which medical personnel would help determine the

actual possible reduction in the single frame dose for which the denoised images are still clinically applicable.

Chapter 6

Conclusion

In this thesis we proposed the ICI-EPL – an adaptive wavelet transform based on the lifting scheme, with a target application in signal denoising. Purpose of the adaptive algorithm was to improve performance of the conventional wavelet transforms about edges in a signal, while still retaining all of their good properties.

The edge preserving nature of the ICI-EPL is achieved by selecting appropriate wavelet basis, independently, for each signal point and on each scale. For smooth signal regions, longer and smoother wavelet basis are used. As we come closer to the edge, more compact wavelets are selected in order to avoid spanning of lifting filters support across the edge.

The edge is considered in a broad sense, and encompasses not only a step edge, but also any other sudden change in signal statistics. The intersection of confidence intervals (ICI) method is used to detect the change in the statistics, and to determine support of the lifting filters, effectively selecting one of the wavelets from a predefined set. We showed that the ICI-EPL is able to efficiently reconstruct both the edges and the smooth regions in the same signal, leading to improved overall denoising performance. In many cases, it is significantly better, while in the other, at least comparable to that of the best performing compared transform.

Denoising efficiency of the ICI-EPL is heavily dependent on the proper choice of the ICI Γ parameter value. Although manual selection can be employed with success, we also proposed a statistical method for automated parameter selection. For the purpose, we consider wavelet coefficients to be realizations of an independent and identically distributed random variable. Distribution of the coefficients varies, based on the analyzed signal and applied transform, with difference being clearly visible at the last decomposition level. The closest model distribution is generalized Gaussian distribution (GGD). The proposed selection method works by finding and comparing actual distribution to the hypothetical GGD the coefficients might come close to taking, in case of a perfect fit of a transform to a given signal. Comparison is carried out using Kolmogorov-Smirnov goodness-of-fit test. The best fitting distribution signals that the respective transform/Γ parameter are to be selected as the optimal ones.

We used the ICI-EPL and its edge preserving nature to improve denoising performance for medical X-ray fluoroscopy image sequences. The application particularities make clear edges in a denoised image an essential requirement. To achieve this, two estimates of denoised images are produced. One is based on converting the image sequence to a 1-D signal and applying the ICI-EPL, before converting it back. The other is obtained by a basic ICI rule applied in the temporal direction. The estimates are fused to form the final denoised image.

The proposed adaptive ICI-EPL transform efficiently combines the statistical ICI method and the lifting scheme wavelet transform realization. As shown in numerous examples, denoising efficiency is greatly improved when compared to conventional wavelet transforms. Most improvement is achieved about edges in a signal, which tend to be most difficult to efficiently reconstruct. The ICI-EPL demonstrated a performance which brings confidence that the proposed transform can be successfully utilized in a variety of denoising applications.

Appendices

Appendix A

Denoising Performance Tables for the ICI-EPL

Denoising performance of the ICI-EPL adaptive lifting scheme was evaluated by denoising 6 different test signals. They are generated as if they were samples of the respective continuous time functions, taken at different sampling frequency. We chose 5 different frequencies to obtain discrete time test signals, whose lengths are from the following set: $\{256, 512, 1024, 2048, 4096\}$. Levels of additive white Gaussian noise added to the signals were: $\sigma_\epsilon \in \{5\%, 10\%, 15\%, 20\%\}$ of total signal magnitude. In Chapter 3.5, the performance was thoroughly examined on two selected signal length/noise combinations. In this Appendix, we present root mean square error (RMSE) tables of denoised signals for all considered signal lengths and noise levels.

Table A.1: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Signal length is 256 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	2.17	2.93	3.11	3.17	3.95	3.92	**1.81**
Bumps	**3.17**	3.20	3.26	3.23	3.59	6.72	3.22
Doppler	3.68	2.93	**2.71**	2.83	2.74	3.82	2.81
HeaviSine	**1.99**	2.05	2.13	2.12	2.22	2.18	2.07
Piece-Polynomial	2.12	2.90	3.04	2.99	3.70	3.47	**1.96**
Piece-Regular	2.86	2.87	2.91	2.87	3.60	3.19	**2.69**

Table A.2: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Signal length is 256 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	4.76	6.22	6.27	6.18	7.18	6.13	**4.45**
Bumps	5.95	**5.79**	6.03	5.94	6.62	8.07	5.89
Doppler	6.48	5.72	**5.35**	5.59	5.66	6.41	5.47
HeaviSine	**3.38**	3.56	3.70	3.65	3.79	4.11	3.66
Piece-Polynomial	4.38	5.36	5.56	5.47	6.04	5.28	**4.19**
Piece-Regular	**5.28**	5.50	5.82	5.72	6.31	5.67	5.38

Table A.3: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Signal length is 256 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	7.15	8.44	8.74	8.64	9.28	8.40	**6.85**
Bumps	8.14	**7.89**	8.20	8.16	8.77	9.13	7.97
Doppler	8.73	7.81	7.78	8.03	8.29	8.74	**7.66**
HeaviSine	**4.78**	5.11	5.30	5.24	5.44	5.59	5.27
Piece-Polynomial	**6.21**	7.23	7.46	7.33	7.98	7.06	6.25
Piece-Regular	**7.34**	7.42	7.87	7.62	8.20	7.73	7.36

Table A.4: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Signal length is 256 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	9.32	10.28	10.72	10.60	11.00	10.64	**9.27**
Bumps	9.76	9.72	10.00	9.99	10.57	9.89	**9.66**
Doppler	10.54	9.67	9.81	10.09	10.42	10.92	**9.62**
HeaviSine	**6.21**	6.69	6.95	6.86	7.13	7.17	6.92
Piece-Polynomial	**7.96**	8.85	9.13	9.02	9.55	8.74	7.98
Piece-Regular	**9.14**	9.17	9.57	9.39	9.84	9.46	9.24

Table A.5: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Signal length is 512 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	1.63	2.34	2.50	2.50	3.30	3.07	**1.27**
Bumps	2.60	**2.40**	2.55	2.67	3.18	4.89	2.53
Doppler	2.98	2.37	**2.13**	2.22	2.26	2.56	2.21
HeaviSine	1.56	1.56	1.62	1.61	1.79	1.63	**1.50**
Piece-Polynomial	1.63	2.35	2.46	2.37	3.10	2.48	**1.30**
Piece-Regular	2.13	2.12	2.16	2.13	2.92	2.27	**1.83**

Table A.6: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Signal length is 512 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	3.61	5.00	5.18	5.11	6.05	4.90	**3.04**
Bumps	5.19	**4.63**	4.95	5.00	5.86	6.65	4.85
Doppler	4.97	4.16	3.91	4.12	4.20	4.65	**3.89**
HeaviSine	**2.48**	2.65	2.73	2.71	2.80	3.04	2.73
Piece-Polynomial	3.40	4.50	4.68	4.56	5.15	4.49	**3.19**
Piece-Regular	3.83	4.08	4.29	4.20	4.95	4.07	**3.81**

Table A.7: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Signal length is 512 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	5.61	7.00	7.24	7.24	7.93	6.55	**4.97**
Bumps	7.30	**6.64**	7.12	7.07	7.90	7.77	6.83
Doppler	6.71	5.74	5.60	5.90	6.02	6.66	**5.39**
HeaviSine	**3.43**	3.64	3.76	3.72	3.85	4.01	3.78
Piece-Polynomial	5.16	5.90	6.05	5.97	6.51	5.90	**4.93**
Piece-Regular	**5.55**	5.69	6.04	5.85	6.50	5.87	**5.56**

Table A.8: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Signal length is 512 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	7.35	8.43	8.75	8.78	9.33	8.15	**6.90**
Bumps	8.99	**8.26**	8.82	8.82	9.44	8.71	8.32
Doppler	8.20	7.22	7.17	7.53	7.89	8.22	**6.92**
HeaviSine	**4.35**	4.67	4.84	4.78	4.95	4.96	4.83
Piece-Polynomial	6.30	6.98	7.13	7.02	7.70	7.05	**6.16**
Piece-Regular	**6.94**	7.03	7.33	7.16	7.61	7.39	7.05

Table A.9: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Signal length is 1024 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	1.22	1.86	2.01	1.99	2.70	2.20	**0.89**
Bumps	1.99	**1.91**	1.99	2.03	2.53	2.25	1.94
Doppler	2.49	1.84	**1.66**	1.76	1.71	2.01	1.69
HeaviSine	1.27	1.14	1.18	1.18	1.41	1.14	**1.09**
Piece-Polynomial	1.21	1.79	1.85	1.84	2.47	1.73	**0.95**
Piece-Regular	1.73	1.75	1.80	1.72	2.32	1.80	**1.41**

Table A.10: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Signal length is 1024 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	2.71	3.88	4.02	3.96	4.85	3.81	**2.25**
Bumps	3.85	**3.67**	3.92	3.90	4.67	4.00	3.73
Doppler	4.12	3.34	3.15	3.30	3.34	3.77	**3.09**
HeaviSine	**1.91**	2.04	2.13	2.10	2.23	2.29	2.08
Piece-Polynomial	2.50	3.38	3.59	3.49	4.27	3.35	**2.22**
Piece-Regular	3.12	3.21	3.47	3.37	4.27	3.20	**2.96**

Table A.11: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Signal length is 1024 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	4.09	5.39	5.70	5.69	6.54	5.04	**3.67**
Bumps	5.71	**5.21**	5.61	5.53	6.36	5.64	5.33
Doppler	5.57	4.62	4.43	4.66	4.79	5.24	**4.36**
HeaviSine	**2.58**	2.77	2.87	2.83	2.93	3.18	2.85
Piece-Polynomial	3.71	4.69	4.97	4.80	5.53	4.76	**3.63**
Piece-Regular	4.50	4.61	5.00	4.84	5.63	4.63	**4.47**

Table A.12: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Signal length is 1024 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	5.42	6.67	7.04	6.99	7.80	6.34	**5.21**
Bumps	7.35	**6.57**	7.12	7.02	7.68	6.83	6.82
Doppler	6.76	5.77	5.59	5.86	6.11	6.53	**5.51**
HeaviSine	**3.24**	3.48	3.60	3.56	3.67	4.00	3.61
Piece-Polynomial	4.86	5.67	5.97	5.85	6.50	5.82	**4.83**
Piece-Regular	**5.72**	5.78	6.20	6.00	6.69	5.86	5.78

Table A.13: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Signal length is 2048 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	0.87	1.46	1.55	1.53	2.14	1.61	**0.67**
Bumps	1.56	**1.39**	1.46	1.46	1.91	1.67	1.42
Doppler	1.89	1.34	1.18	1.26	**1.17**	1.54	1.22
HeaviSine	0.98	0.84	0.87	0.86	1.07	0.94	**0.79**
Piece-Polynomial	0.91	1.34	1.42	1.40	1.93	1.35	**0.73**
Piece-Regular	1.27	1.29	1.32	1.28	1.76	1.31	**1.05**

Table A.14: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Signal length is 2048 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	1.99	3.04	3.19	3.18	3.98	2.92	**1.55**
Bumps	2.80	**2.64**	2.81	2.81	3.58	2.97	2.68
Doppler	3.11	2.43	2.25	2.36	2.35	2.88	**2.22**
HeaviSine	**1.60**	**1.60**	1.69	1.65	1.84	1.86	1.64
Piece-Polynomial	1.87	2.60	2.77	2.72	3.46	2.58	**1.69**
Piece-Regular	2.39	2.48	2.63	2.60	3.36	2.41	**2.09**

Table A.15: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Signal length is 2048 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	3.13	4.31	4.58	4.52	5.38	4.18	**2.80**
Bumps	4.15	**3.86**	4.18	4.16	4.97	4.15	3.98
Doppler	4.15	3.37	3.19	3.34	3.37	3.95	**3.12**
HeaviSine	**2.05**	2.20	2.29	2.24	2.36	2.66	2.26
Piece-Polynomial	**2.84**	3.69	3.95	3.85	4.61	3.74	2.86
Piece-Regular	**3.41**	3.58	3.82	3.71	4.55	3.56	**3.41**

Table A.16: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Signal length is 2048 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	4.20	5.44	5.75	5.69	6.58	5.33	**4.13**
Bumps	5.50	**4.98**	5.43	5.32	6.27	5.24	5.19
Doppler	5.11	4.29	4.07	4.25	4.35	5.02	**4.01**
HeaviSine	**2.52**	2.70	2.80	2.77	2.86	3.41	2.77
Piece-Polynomial	**3.69**	4.57	4.89	4.78	5.50	4.67	3.89
Piece-Regular	**4.42**	4.56	4.90	4.74	5.50	4.61	4.61

Table A.17: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Signal length is 4096 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	0.69	1.15	1.24	1.22	1.70	1.25	**0.50**
Bumps	1.18	**1.04**	1.13	1.11	1.46	1.28	1.08
Doppler	1.46	1.04	0.89	0.95	**0.87**	1.20	0.93
HeaviSine	0.74	0.64	0.67	0.67	0.86	0.85	**0.58**
Piece-Polynomial	0.71	1.04	1.13	1.11	1.53	1.10	**0.58**
Piece-Regular	0.98	0.99	1.01	0.98	1.34	1.03	**0.80**

Table A.18: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Signal length is 4096 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	1.56	2.40	2.57	2.53	3.22	2.34	**1.19**
Bumps	2.11	**1.99**	2.15	2.13	2.73	2.34	2.05
Doppler	2.44	1.85	1.68	1.76	1.70	2.27	**1.65**
HeaviSine	1.35	**1.24**	1.30	1.27	1.50	1.67	1.28
Piece-Polynomial	1.43	2.00	2.15	2.11	2.75	2.15	**1.29**
Piece-Regular	1.81	1.84	1.99	1.95	2.57	1.93	**1.55**

Table A.19: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Signal length is 4096 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	2.44	3.43	3.69	3.63	4.42	3.45	**2.09**
Bumps	3.03	**2.88**	3.17	3.13	3.92	3.31	2.99
Doppler	3.33	2.65	2.51	2.59	2.58	3.23	**2.46**
HeaviSine	**1.67**	1.74	1.84	1.78	1.96	2.41	1.82
Piece-Polynomial	**2.13**	2.86	3.10	3.01	3.76	3.09	2.18
Piece-Regular	2.55	2.69	2.91	2.83	3.59	2.86	**2.50**

Table A.20: RMSE values of denoised signals for additive Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Signal length is 4096 points. RMSE multiplied by a 10^2

	Haar	Bior2.2	Bior4.4	Db3	Db9	ICI	ICI-EPL
Blocks	3.25	4.35	4.68	4.60	5.44	4.50	**3.16**
Bumps	4.02	**3.74**	4.15	4.07	4.94	4.22	3.96
Doppler	4.09	3.36	3.20	3.31	3.36	4.11	**3.14**
HeaviSine	**2.00**	2.16	2.27	2.23	2.32	3.10	2.26
Piece-Polynomial	**2.76**	3.65	3.96	3.83	4.57	3.92	3.26
Piece-Regular	**3.28**	3.46	3.77	3.60	4.45	3.74	3.52

Appendix B

Parameter Γ Selection Method Performance Tables

Two statistical methods for automated selection of the ICI-EPL Γ parameter value were proposed in Chapter 4. Their efficiency was analyzed and discussed on a several selected test case examples.

In this Appendix, we present the performance measures, in terms of the root mean square error (RMSE) for all the considered test cases.

Table B.1: RMSE values of denoised signals. Signal length is 256. Level of additive Gaussian noise: $\sigma_\epsilon = 5\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	2.17	3.92	1.81	1.89	2.07	2.37	**1.84**	**1.84**
Bumps	3.17	6.72	3.22	3.60	**3.34**	3.61	3.85	3.62
Doppler	2.71	3.82	2.81	3.07	3.41	3.18	2.85	**2.83**
HeaviSine	1.99	2.18	2.07	2.18	**2.16**	2.18	2.35	2.18
Piece-Polynomial	2.12	3.47	1.96	2.04	2.15	2.32	**2.02**	**2.02**
Piece-Regular	2.86	3.19	2.69	2.99	2.95	**2.88**	3.50	3.55

Table B.2: RMSE values of denoised signals. Signal length is 256. Level of additive Gaussian noise: $\sigma_\epsilon = 10\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	4.76	6.13	4.45	4.73	5.08	5.02	**4.58**	**4.58**
Bumps	5.79	8.07	5.89	**6.14**	6.14	6.22	6.21	6.28
Doppler	5.35	6.41	5.47	**5.55**	6.41	5.72	5.77	5.77
HeaviSine	3.38	4.11	3.66	3.76	3.75	**3.75**	3.93	3.78
Piece-Polynomial	4.38	5.28	4.19	**4.45**	4.78	4.69	4.51	4.50
Piece-Regular	5.28	5.67	5.38	5.90	**5.71**	5.83	6.21	6.12

Table B.3: RMSE values of denoised signals. Signal length is 256. Level of additive Gaussian noise: $\sigma_\epsilon = 15\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	7.15	8.40	6.85	**7.29**	7.41	7.35	7.34	7.34
Bumps	7.89	9.13	7.97	8.30	**8.21**	8.32	8.24	8.35
Doppler	7.78	8.74	7.66	7.72	8.13	**7.72**	8.24	8.24
HeaviSine	4.78	5.59	5.27	5.36	**5.32**	5.37	5.38	5.37
Piece-Polynomial	6.21	7.06	6.25	6.69	**6.50**	6.59	6.60	6.68
Piece-Regular	7.34	7.73	7.36	7.81	**7.74**	7.78	7.84	7.87

Table B.4: RMSE values of denoised signals. Signal length is 256. Level of additive Gaussian noise: $\sigma_\epsilon = 20\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	9.32	10.64	9.27	9.83	**9.68**	9.81	10.21	10.23
Bumps	9.72	9.89	9.66	10.07	10.03	10.10	10.03	**10.01**
Doppler	9.67	10.92	9.62	10.11	**9.91**	10.24	10.07	9.96
HeaviSine	6.21	7.17	6.92	6.99	7.03	6.98	7.02	**6.97**
Piece-Polynomial	7.96	8.74	7.98	8.41	8.33	**8.28**	8.38	8.59
Piece-Regular	9.14	9.46	9.24	9.57	9.53	**9.53**	9.61	9.53

Table B.5: RMSE values of denoised signals. Signal length is 512. Level of additive Gaussian noise: $\sigma_\epsilon = 5\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	1.63	3.07	1.27	1.32	1.45	1.59	**1.30**	**1.30**
Bumps	2.40	4.89	2.53	2.71	**2.60**	2.86	3.19	3.19
Doppler	2.13	2.56	2.21	2.30	2.33	2.31	**2.24**	**2.24**
HeaviSine	1.56	1.63	1.50	1.64	**1.59**	1.61	1.59	1.59
Piece-Polynomial	1.63	2.48	1.30	**1.35**	1.37	1.38	1.39	1.39
Piece-Regular	2.12	2.27	1.83	2.20	2.16	**2.02**	2.59	2.59

Table B.6: RMSE values of denoised signals. Signal length is 512. Level of additive Gaussian noise: $\sigma_\epsilon = 10\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	3.61	4.90	3.04	3.42	3.43	3.83	**3.37**	**3.37**
Bumps	4.63	6.65	4.85	5.23	**4.94**	5.51	5.40	5.40
Doppler	3.91	4.65	3.89	3.91	3.91	**3.91**	3.92	3.92
HeaviSine	2.48	3.04	2.73	2.80	2.80	**2.80**	2.81	2.81
Piece-Polynomial	3.40	4.49	3.19	**3.68**	3.81	3.86	4.35	4.35
Piece-Regular	3.83	4.07	3.81	4.17	4.21	**4.13**	4.44	4.44

Table B.7: RMSE values of denoised signals. Signal length is 512. Level of additive Gaussian noise: $\sigma_\epsilon = 15\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	5.61	6.55	4.97	5.39	5.51	5.57	**5.27**	**5.27**
Bumps	6.64	7.77	6.83	7.02	**6.95**	7.13	7.22	7.22
Doppler	5.60	6.66	5.39	5.47	5.50	**5.45**	5.69	5.69
HeaviSine	3.43	4.01	3.78	3.82	3.82	3.83	**3.80**	**3.80**
Piece-Polynomial	5.16	5.90	4.93	5.44	5.43	**5.40**	5.56	5.56
Piece-Regular	5.55	5.87	5.56	**5.83**	5.84	5.85	6.06	6.06

Table B.8: RMSE values of denoised signals. Signal length is 512. Level of additive Gaussian noise: $\sigma_\epsilon = 20\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	7.35	8.15	6.90	7.68	7.56	7.62	**7.31**	**7.31**
Bumps	8.26	8.71	8.32	8.69	8.66	8.65	**8.63**	**8.63**
Doppler	7.17	8.22	6.92	7.14	7.04	**7.02**	7.46	7.46
HeaviSine	4.35	4.96	4.83	4.90	**4.89**	4.90	4.89	4.89
Piece-Polynomial	6.30	7.05	6.16	6.41	6.59	6.47	**6.41**	**6.41**
Piece-Regular	6.94	7.39	7.05	7.24	**7.23**	7.25	7.32	7.32

Table B.9: RMSE values of denoised signals. Signal length is 1024. Level of additive Gaussian noise: $\sigma_\epsilon = 5\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	1.22	2.20	0.89	0.99	**0.94**	1.47	0.96	0.96
Bumps	1.91	2.25	1.94	1.98	**1.98**	2.42	2.86	2.86
Doppler	1.66	2.01	1.69	1.75	1.70	1.73	**1.70**	**1.70**
HeaviSine	1.14	1.14	1.09	1.17	1.15	1.18	**1.13**	**1.13**
Piece-Polynomial	1.21	1.73	0.95	1.02	**1.01**	1.02	1.06	1.06
Piece-Regular	1.72	1.80	1.41	1.50	1.92	**1.50**	1.52	1.52

Table B.10: RMSE values of denoised signals. Signal length is 1024. Level of additive Gaussian noise: $\sigma_\epsilon = 10\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	2.71	3.81	2.25	2.55	**2.32**	2.61	2.43	2.43
Bumps	3.67	4.00	3.73	3.84	**3.82**	4.33	4.25	4.25
Doppler	3.15	3.77	3.09	3.11	3.14	**3.11**	3.11	3.11
HeaviSine	1.91	2.29	2.08	2.18	2.18	2.19	**2.15**	**2.15**
Piece-Polynomial	2.50	3.35	2.22	2.48	**2.45**	2.52	3.36	3.36
Piece-Regular	3.12	3.20	2.96	**3.12**	3.27	3.16	3.32	3.32

Table B.11: RMSE values of denoised signals. Signal length is 1024. Level of additive Gaussian noise: $\sigma_\epsilon = 15\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	4.09	5.04	3.67	4.06	**3.94**	4.15	4.03	4.03
Bumps	5.21	5.64	5.33	5.45	**5.43**	5.61	5.75	5.75
Doppler	4.43	5.24	4.36	4.37	4.53	4.37	**4.37**	**4.37**
HeaviSine	2.58	3.18	2.85	**2.93**	2.93	2.93	2.93	2.93
Piece-Polynomial	3.71	4.76	3.63	4.41	4.41	**4.39**	4.53	4.53
Piece-Regular	4.50	4.63	4.47	4.63	4.72	4.65	**4.58**	**4.58**

Table B.12: RMSE values of denoised signals. Signal length is 1024. Level of additive Gaussian noise: $\sigma_\epsilon = 20\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	5.42	6.34	5.21	5.64	5.50	5.67	**5.44**	**5.44**
Bumps	6.57	6.83	6.82	**6.88**	6.90	6.88	7.05	7.05
Doppler	5.59	6.53	5.51	**5.53**	5.54	5.53	5.54	5.54
HeaviSine	3.24	4.00	3.61	3.66	**3.66**	3.66	3.67	3.67
Piece-Polynomial	4.86	5.82	4.83	**5.41**	5.43	5.43	5.44	5.44
Piece-Regular	5.72	5.86	5.78	6.04	6.06	**6.03**	6.04	6.04

Table B.13: RMSE values of denoised signals. Signal length is 2048. Level of additive Gaussian noise: $\sigma_\epsilon = 5\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	0.87	1.61	0.67	0.71	0.72	0.94	**0.69**	**0.69**
Bumps	1.39	1.67	1.42	1.54	**1.45**	1.92	2.33	2.33
Doppler	1.17	1.54	1.22	1.28	**1.22**	1.23	1.22	1.22
HeaviSine	0.84	0.94	0.79	0.85	0.87	0.88	**0.83**	**0.83**
Piece-Polynomial	0.91	1.35	0.73	0.77	0.77	**0.77**	0.80	0.80
Piece-Regular	1.27	1.31	1.05	1.11	1.40	**1.08**	1.16	1.16

Table B.14: RMSE values of denoised signals. Signal length is 2048. Level of additive Gaussian noise: $\sigma_\epsilon = 10\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	1.99	2.92	1.55	1.75	**1.62**	2.07	1.65	1.65
Bumps	2.64	2.97	2.68	**2.75**	2.76	3.34	3.31	3.31
Doppler	2.25	2.88	2.22	2.23	2.23	**2.23**	2.23	2.23
HeaviSine	1.60	1.86	1.64	1.73	**1.72**	1.73	1.75	1.75
Piece-Polynomial	1.87	2.58	1.69	1.84	**1.83**	1.88	2.54	2.54
Piece-Regular	2.39	2.41	2.09	2.30	2.56	**2.24**	2.28	2.28

Table B.15: RMSE values of denoised signals. Signal length is 2048. Level of additive Gaussian noise: $\sigma_\epsilon = 15\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	3.13	4.18	2.80	3.31	**3.15**	3.42	3.24	3.24
Bumps	3.86	4.15	3.98	4.04	**4.04**	4.26	4.44	4.44
Doppler	3.19	3.95	3.12	3.14	3.14	3.14	**3.14**	**3.14**
HeaviSine	2.05	2.66	2.26	2.36	**2.36**	2.36	2.37	2.37
Piece-Polynomial	2.84	3.74	2.86	3.68	3.67	**3.67**	3.76	3.76
Piece-Regular	3.41	3.56	3.41	3.59	3.69	3.60	**3.56**	**3.56**

Table B.16: RMSE values of denoised signals. Signal length is 2048. Level of additive Gaussian noise: $\sigma_\epsilon = 20\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	4.20	5.33	4.13	4.44	4.44	4.58	**4.36**	**4.36**
Bumps	4.98	5.24	5.19	5.26	**5.24**	5.30	5.50	5.50
Doppler	4.07	5.02	4.01	4.03	4.02	**4.02**	4.03	4.03
HeaviSine	2.52	3.41	2.77	**2.88**	2.88	2.88	2.89	2.89
Piece-Polynomial	3.69	4.67	3.89	4.62	4.60	4.59	**4.56**	**4.56**
Piece-Regular	4.42	4.61	4.61	4.81	4.79	4.78	**4.72**	**4.72**

Table B.17: RMSE values of denoised signals. Signal length is 4096. Level of additive Gaussian noise: $\sigma_\epsilon = 5\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	0.69	1.25	0.50	0.54	0.57	0.60	**0.52**	**0.52**
Bumps	1.04	1.28	1.08	1.15	**1.10**	1.57	1.68	1.68
Doppler	0.87	1.20	0.93	0.96	**0.93**	0.95	0.93	0.93
HeaviSine	0.64	0.85	0.58	0.66	0.64	0.64	**0.60**	**0.60**
Piece-Polynomial	0.71	1.10	0.58	0.61	**0.61**	0.61	0.64	0.64
Piece-Regular	0.98	1.03	0.80	0.85	1.12	**0.82**	0.85	0.85

Table B.18: RMSE values of denoised signals. Signal length is 4096. Level of additive Gaussian noise: $\sigma_\epsilon = 10\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	1.56	2.34	1.19	1.30	**1.27**	1.44	1.30	1.30
Bumps	1.99	2.34	2.05	2.10	**2.10**	2.24	2.49	2.49
Doppler	1.68	2.27	1.65	1.66	**1.66**	1.66	1.67	1.67
HeaviSine	1.24	1.67	1.28	1.35	1.35	1.36	**1.33**	**1.33**
Piece-Polynomial	1.43	2.15	1.29	**1.37**	**1.37**	1.38	1.99	1.99
Piece-Regular	1.81	1.93	1.55	1.70	1.87	**1.68**	1.71	1.71

Table B.19: RMSE values of denoised signals. Signal length is 4096. Level of additive Gaussian noise: $\sigma_\epsilon = 15\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	2.44	3.45	2.09	2.49	**2.39**	2.55	2.46	2.46
Bumps	2.88	3.31	2.99	**3.03**	3.03	3.04	3.38	3.38
Doppler	2.51	3.23	2.46	2.47	**2.46**	2.47	2.47	2.47
HeaviSine	1.67	2.41	1.82	**1.91**	1.91	1.91	1.95	1.95
Piece-Polynomial	2.13	3.09	2.18	2.90	**2.89**	2.90	2.90	2.90
Piece-Regular	2.55	2.86	2.50	2.65	2.74	2.67	**2.57**	**2.57**

Table B.20: RMSE values of denoised signals. Signal length is 4096. Level of additive Gaussian noise: $\sigma_\epsilon = 20\%$ of total signal magnitude. RMSE multiplied by a 10^2

	Optimal fixed wavelet	ICI	ICI-EPL					
			Γ_o	Γ^*_{GGD}			Γ^*_k	Γ^*_{km}
				KS	χ^2	CM		
Blocks	3.25	4.50	3.16	3.56	3.53	3.63	**3.48**	**3.48**
Bumps	3.74	4.22	3.96	3.99	3.99	**3.98**	4.30	4.30
Doppler	3.20	4.11	3.14	3.16	**3.15**	3.15	3.17	3.17
HeaviSine	2.00	3.10	2.26	**2.34**	2.34	2.35	2.36	2.36
Piece-Polynomial	2.76	3.92	3.26	3.86	3.86	3.86	**3.78**	**3.78**
Piece-Regular	3.28	3.74	3.52	**3.66**	3.66	3.68	3.67	3.67

Appendix C

Parameter Γ Selection Method Performance Plots

A statistical method for automated selection of the ICI-EPL Γ parameter value was proposed in Chapter 4.3.2. The method is based on modeling the distribution of wavelet coefficients at the last decomposition level. It was empirically devised and its performance was briefly investigated by presenting a few performance examples only.

In this Appendix, we show all the performance plots of the proposed method for the *Blocks*, *Bumps*, *Doppler*, *HeviSine*, *Piece-Polynomial* and *Piece-Regular* signals. Signal lengths are $\in \{256, 512, 1024, 2048, 4096\}$ samples. Noise corruption was simulated by superimposing the Gaussian white noise to the signals. Noise levels considered are equal to $5\%, 10\%, 15\%$ and 20% of total signal magnitude.

Figures depict the performance of denoising algorithm, for all three goodness-of-fit tests, in terms of the ratio:

$$RMSE_{\Gamma^*_{GGD}} / \min(RMSE),$$

plotted against the GGD shape parameter c. To eliminate influence of particularities of a given noise realization, the performance was averaged over 10 different noise realizations. For more details about the method and discussion of results, see Chapter 4.

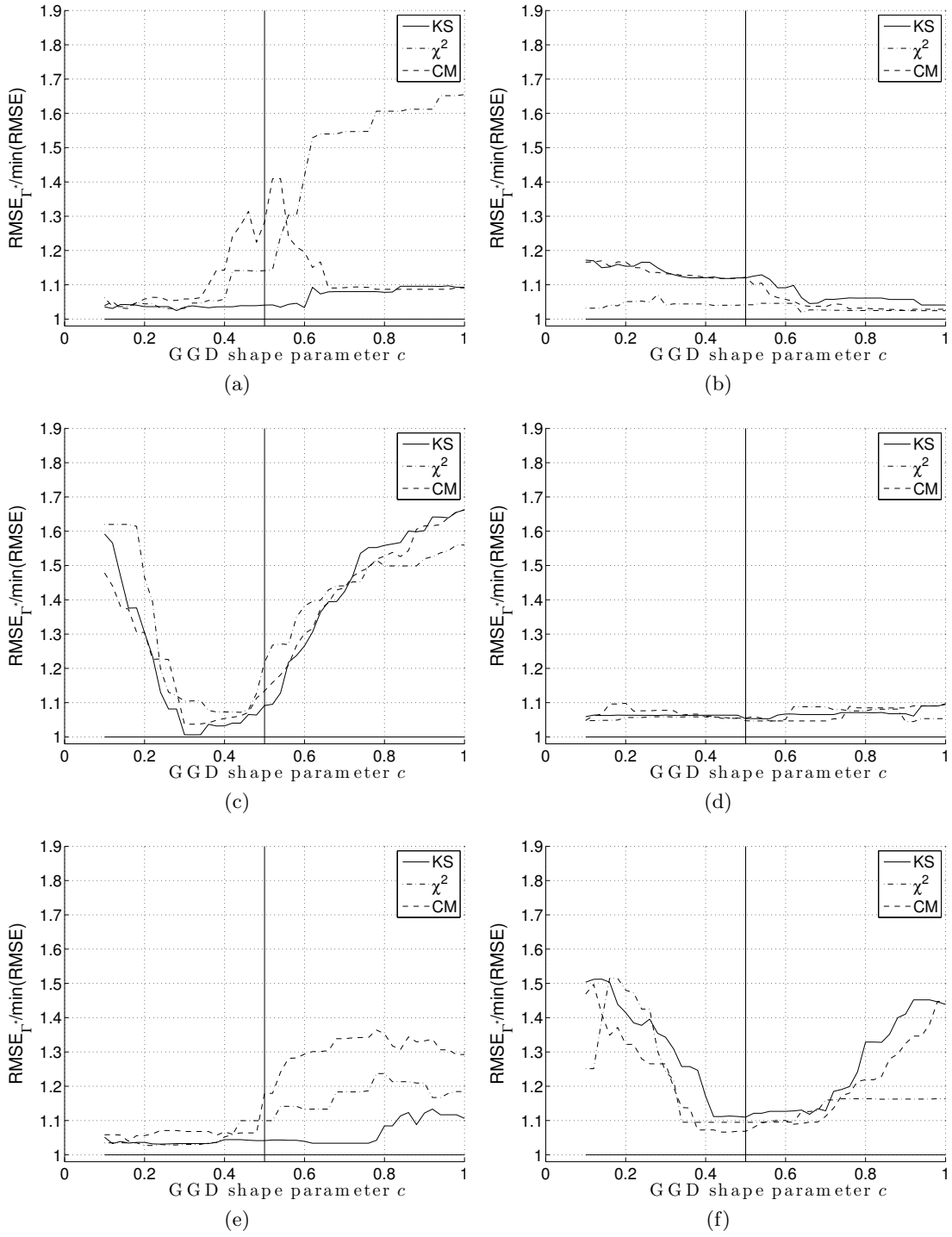


Figure C.1: Performance of the Γ selection method for 256 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

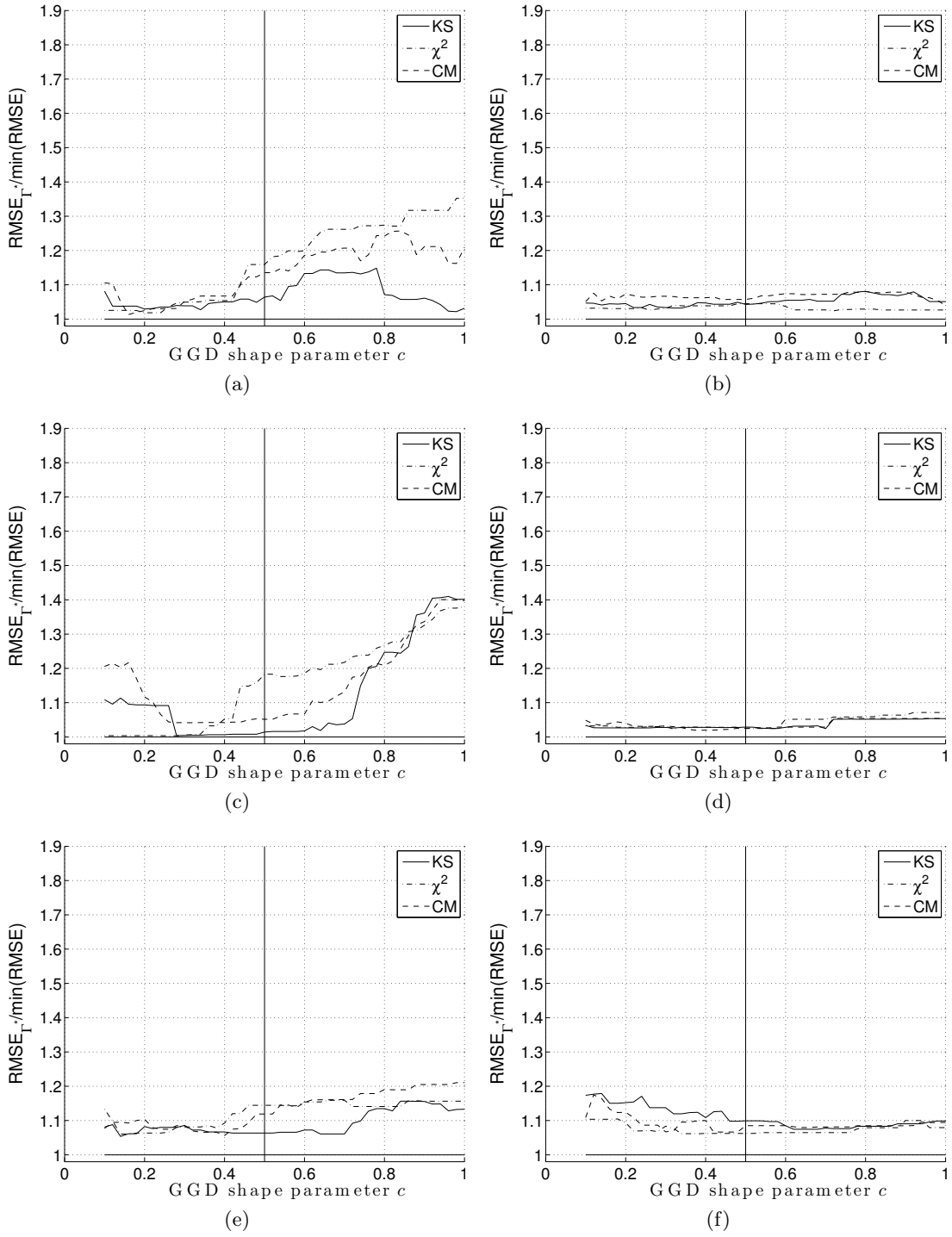


Figure C.2: Performance of the Γ selection method for 256 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

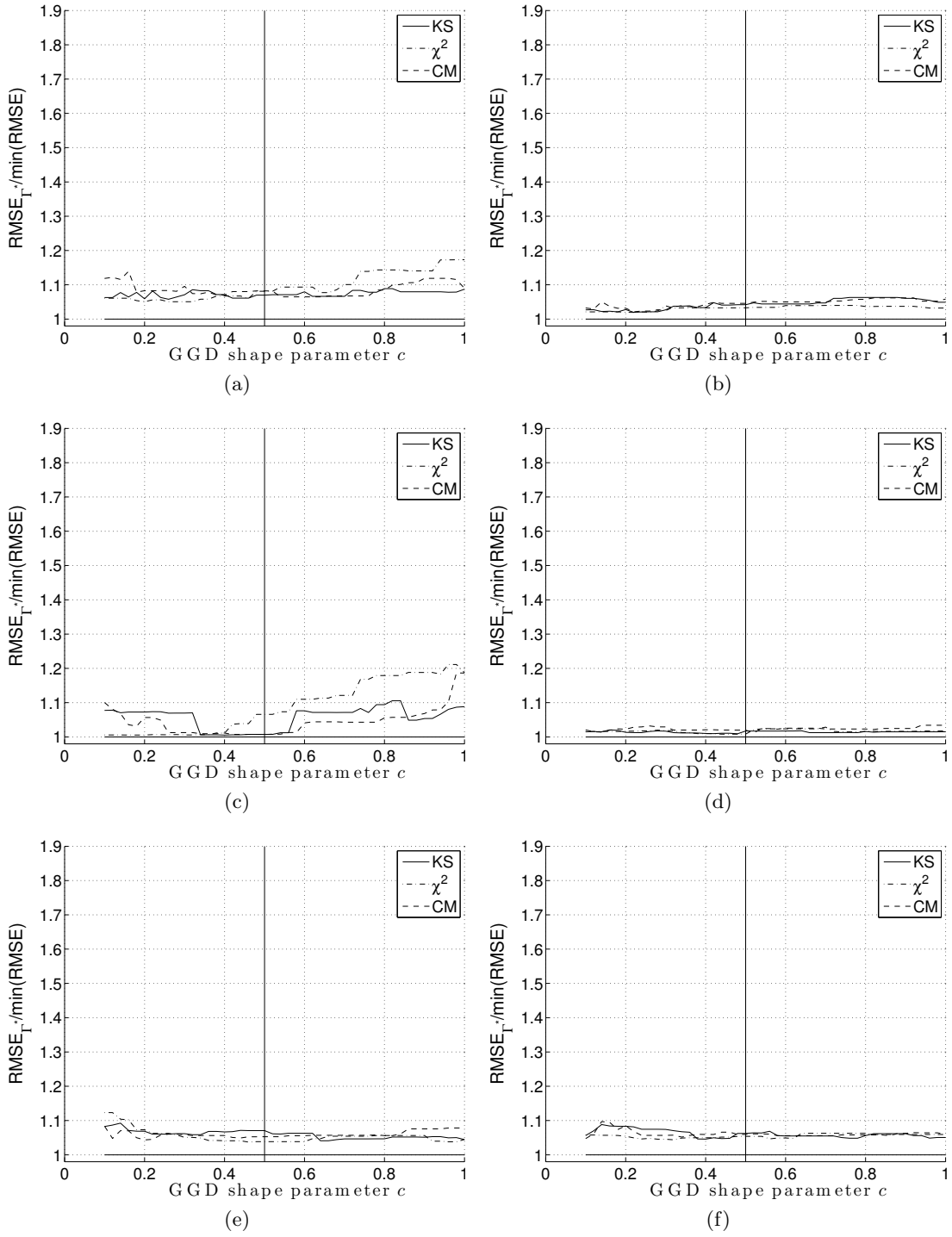


Figure C.3: Performance of the Γ selection method for 256 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

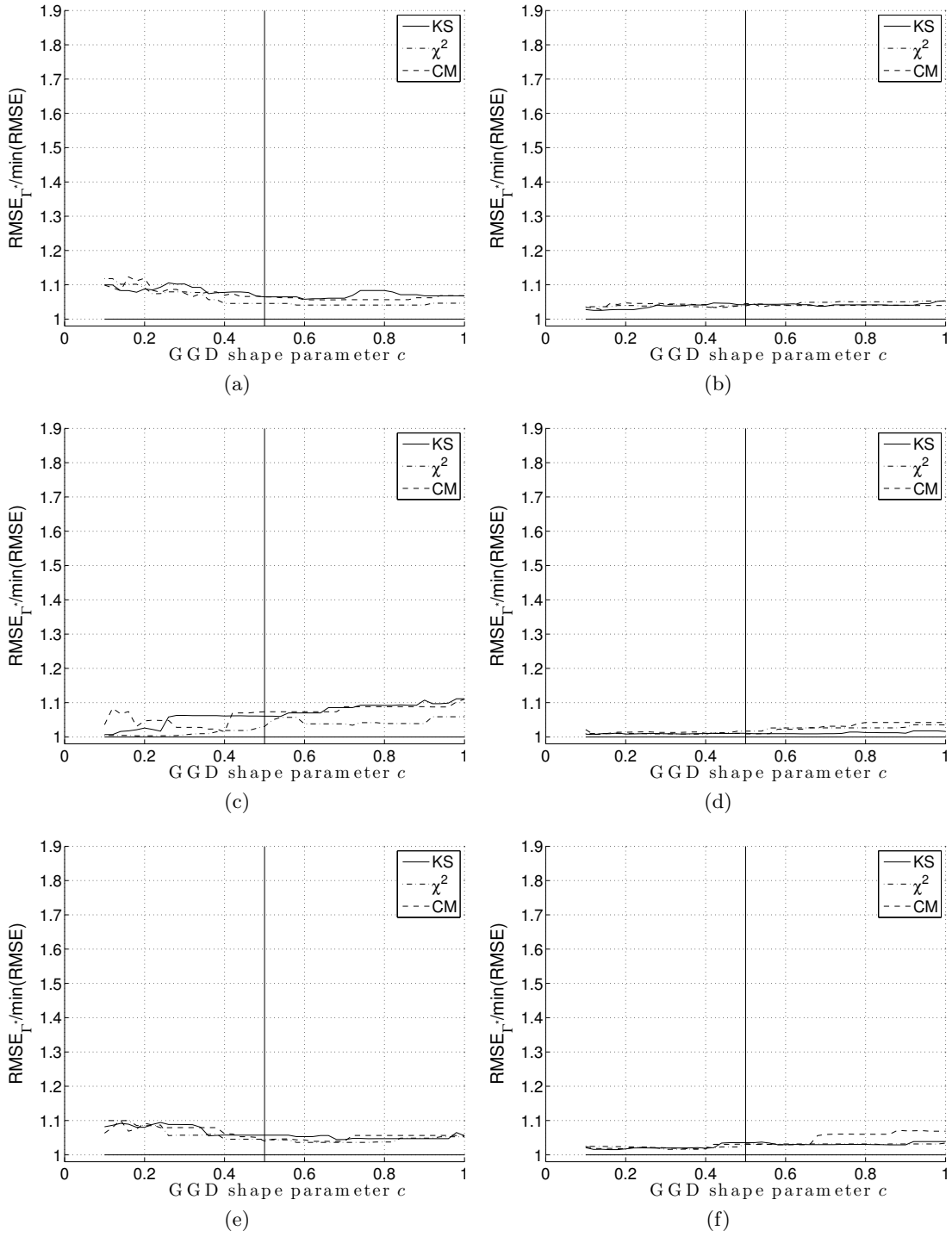


Figure C.4: Performance of the Γ selection method for 256 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

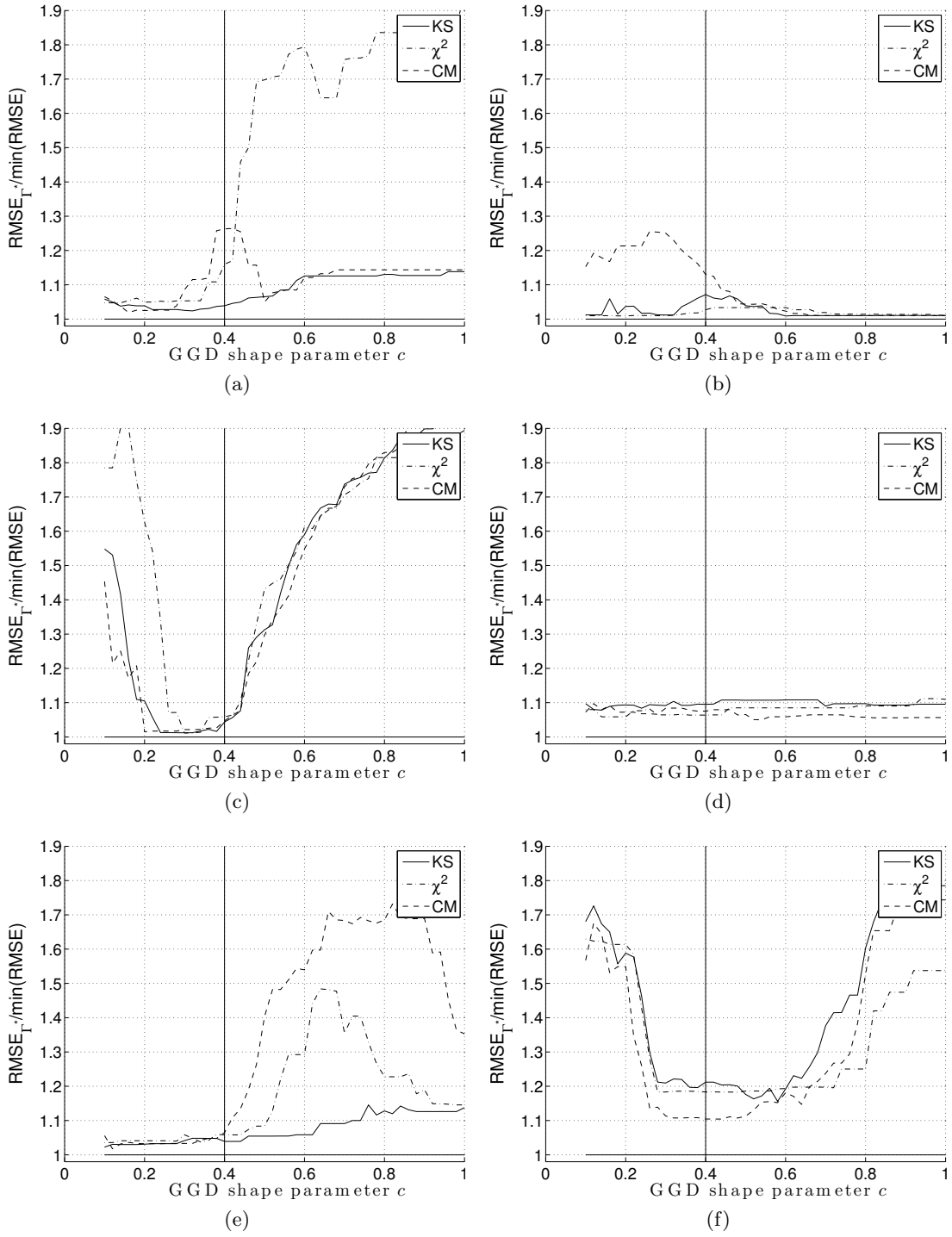


Figure C.5: Performance of the Γ selection method for 512 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

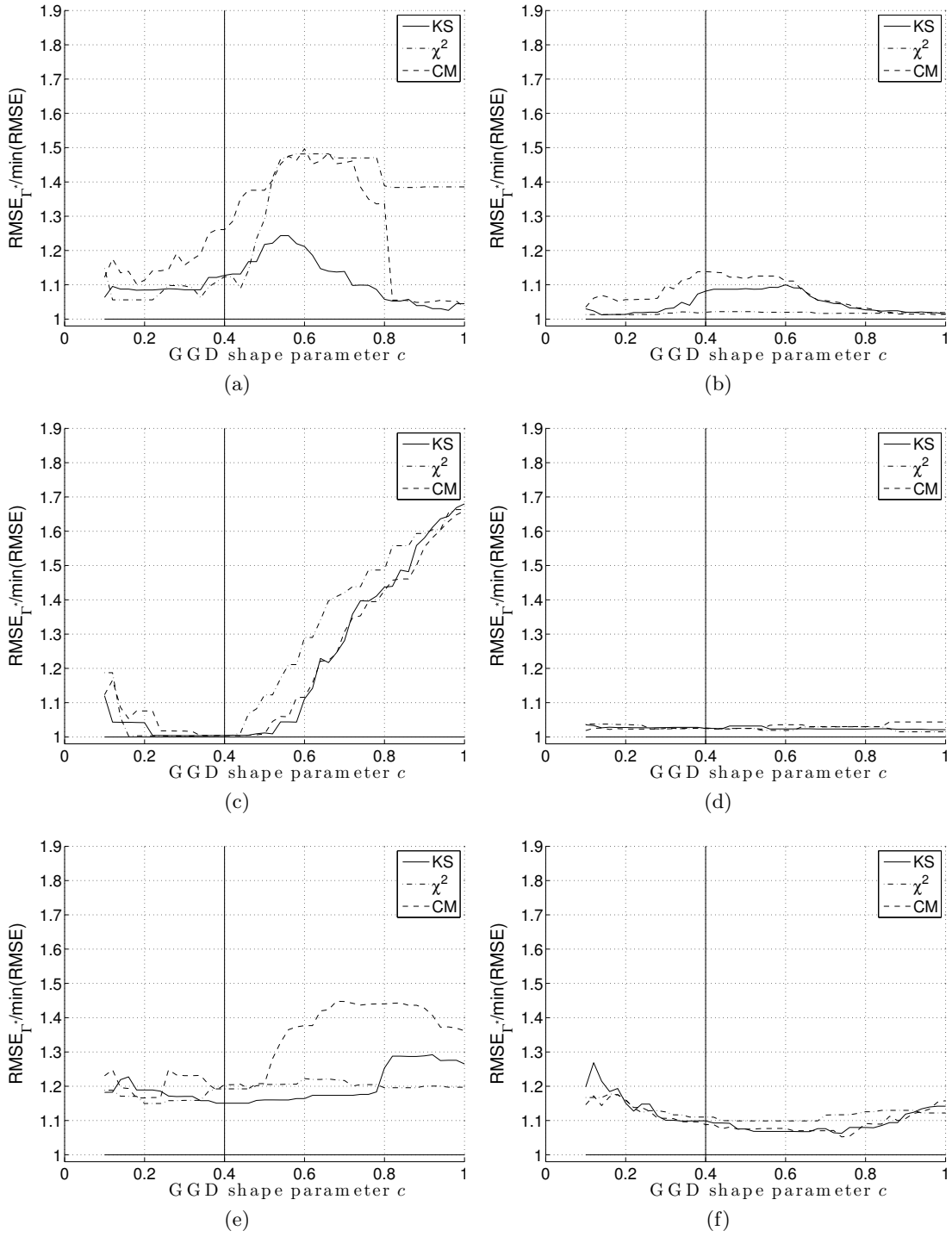


Figure C.6: Performance of the Γ selection method for 512 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

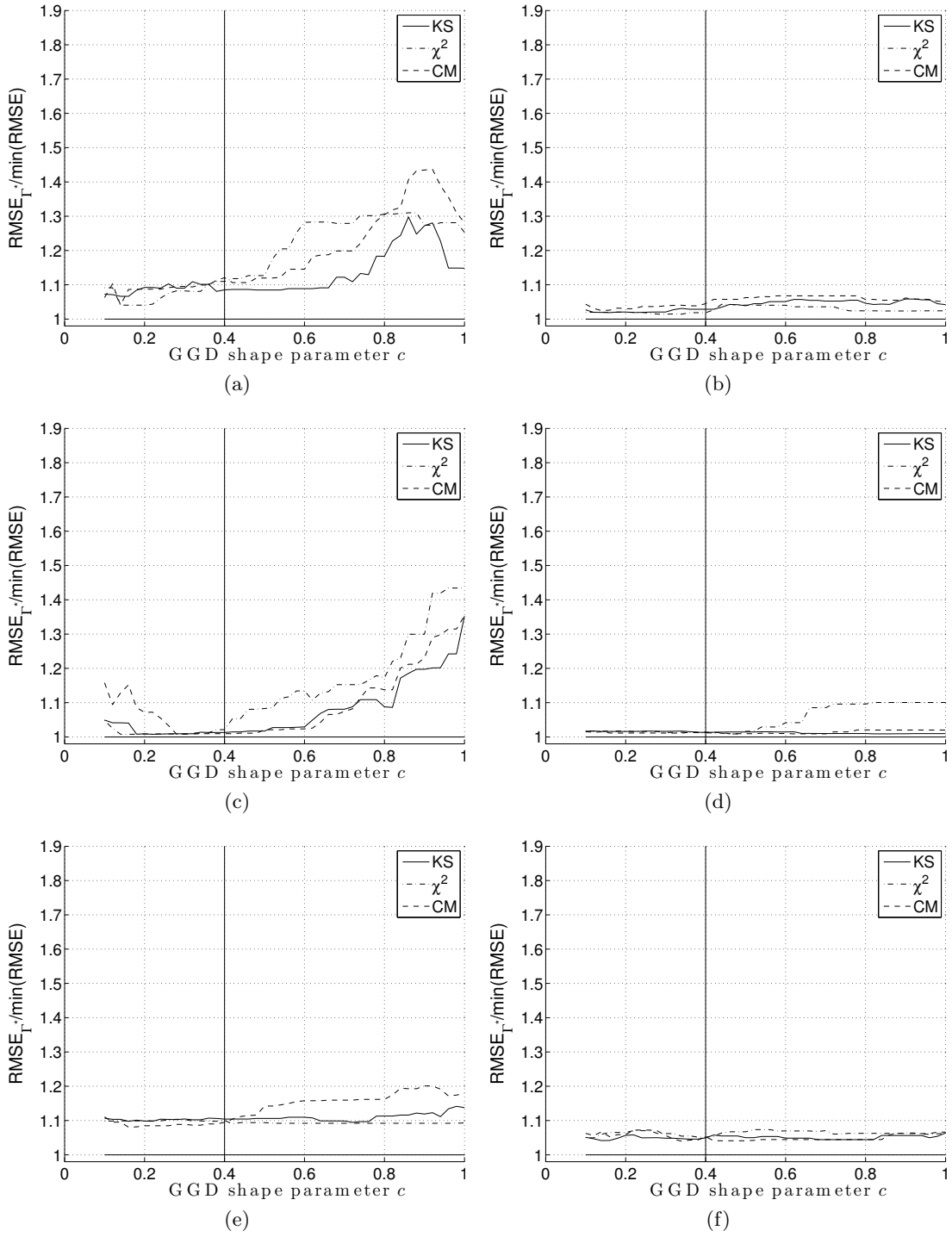


Figure C.7: Performance of the Γ selection method for 512 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

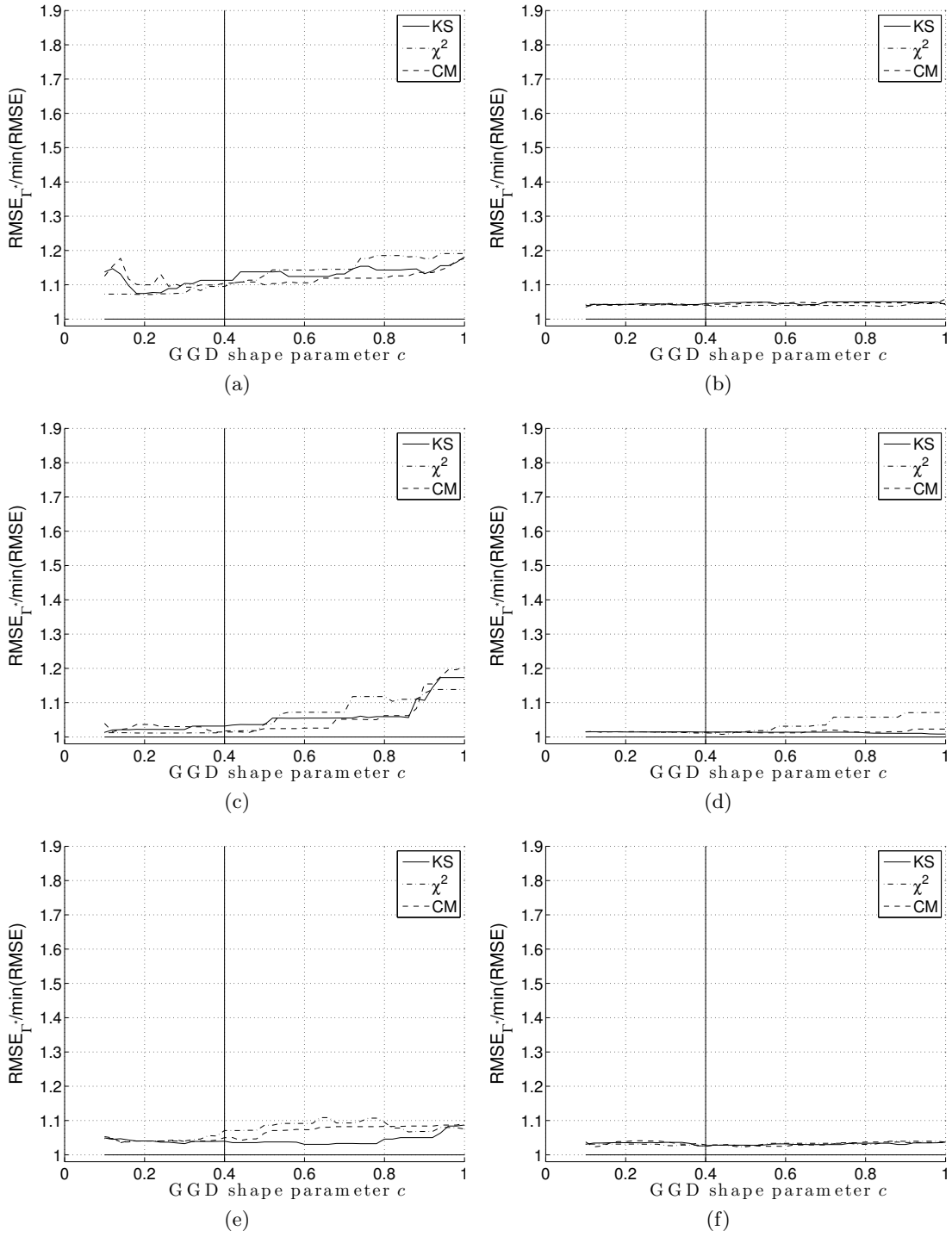


Figure C.8: Performance of the Γ selection method for 512 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

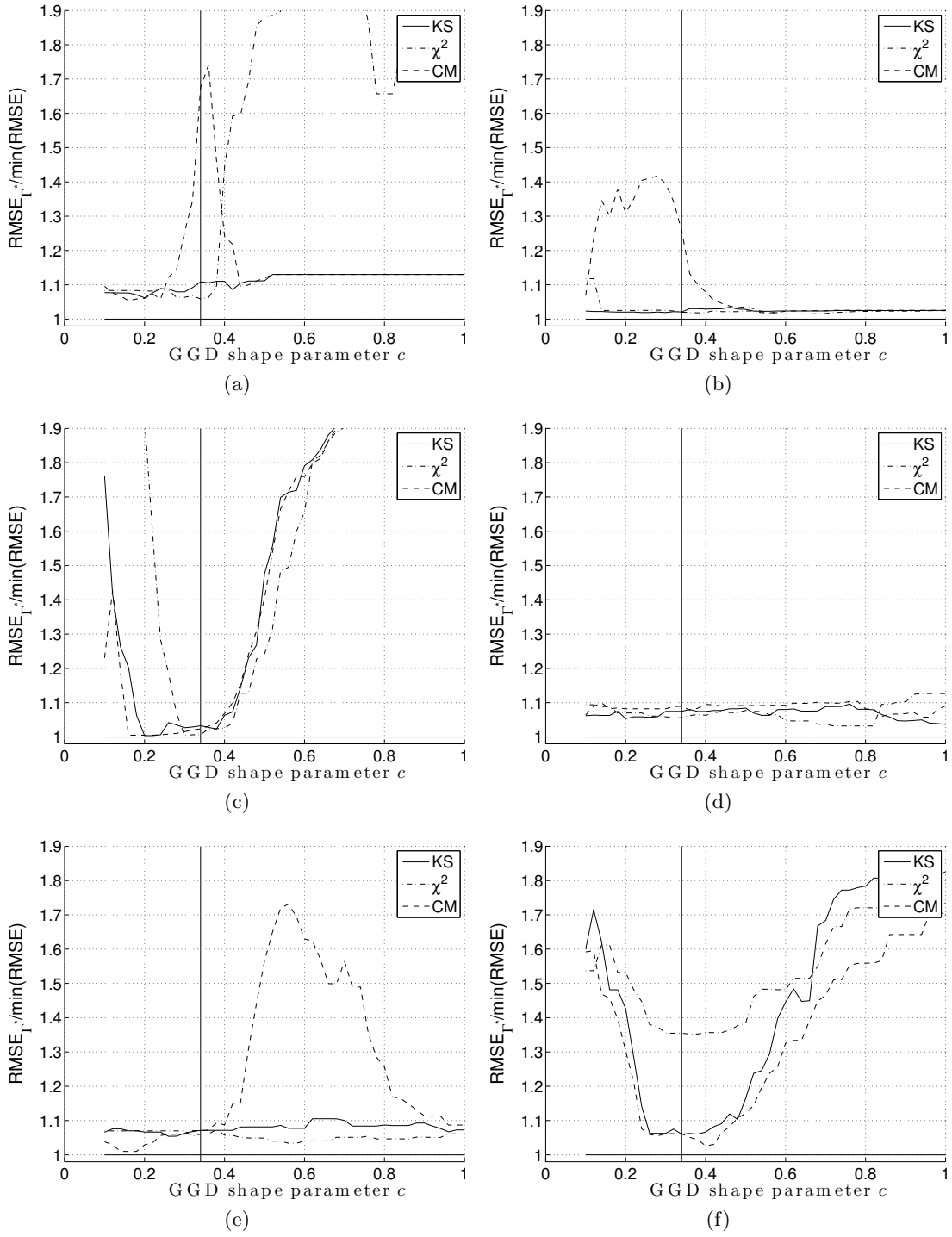


Figure C.9: Performance of the Γ selection method for 1024 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

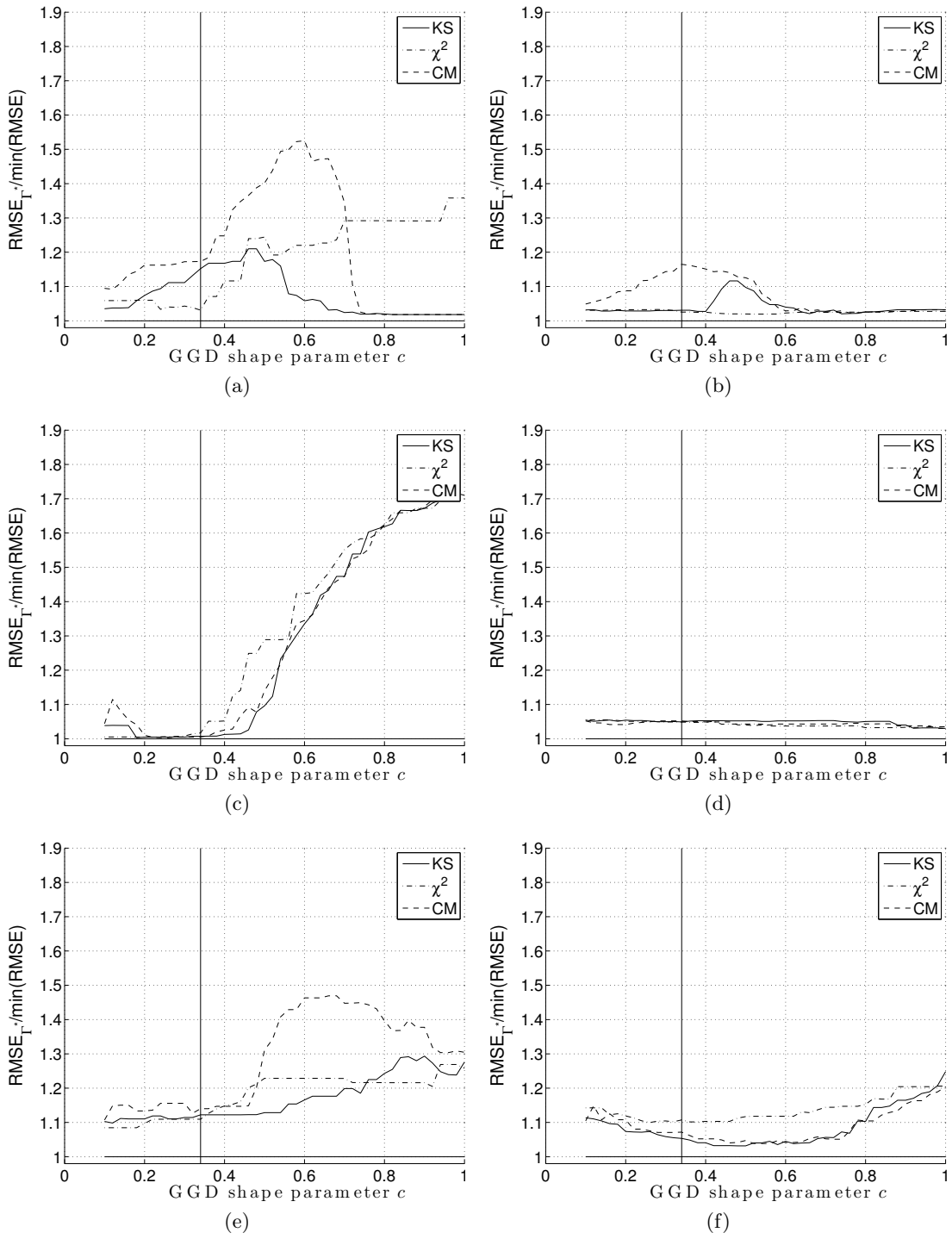


Figure C.10: Performance of the Γ selection method for 1024 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

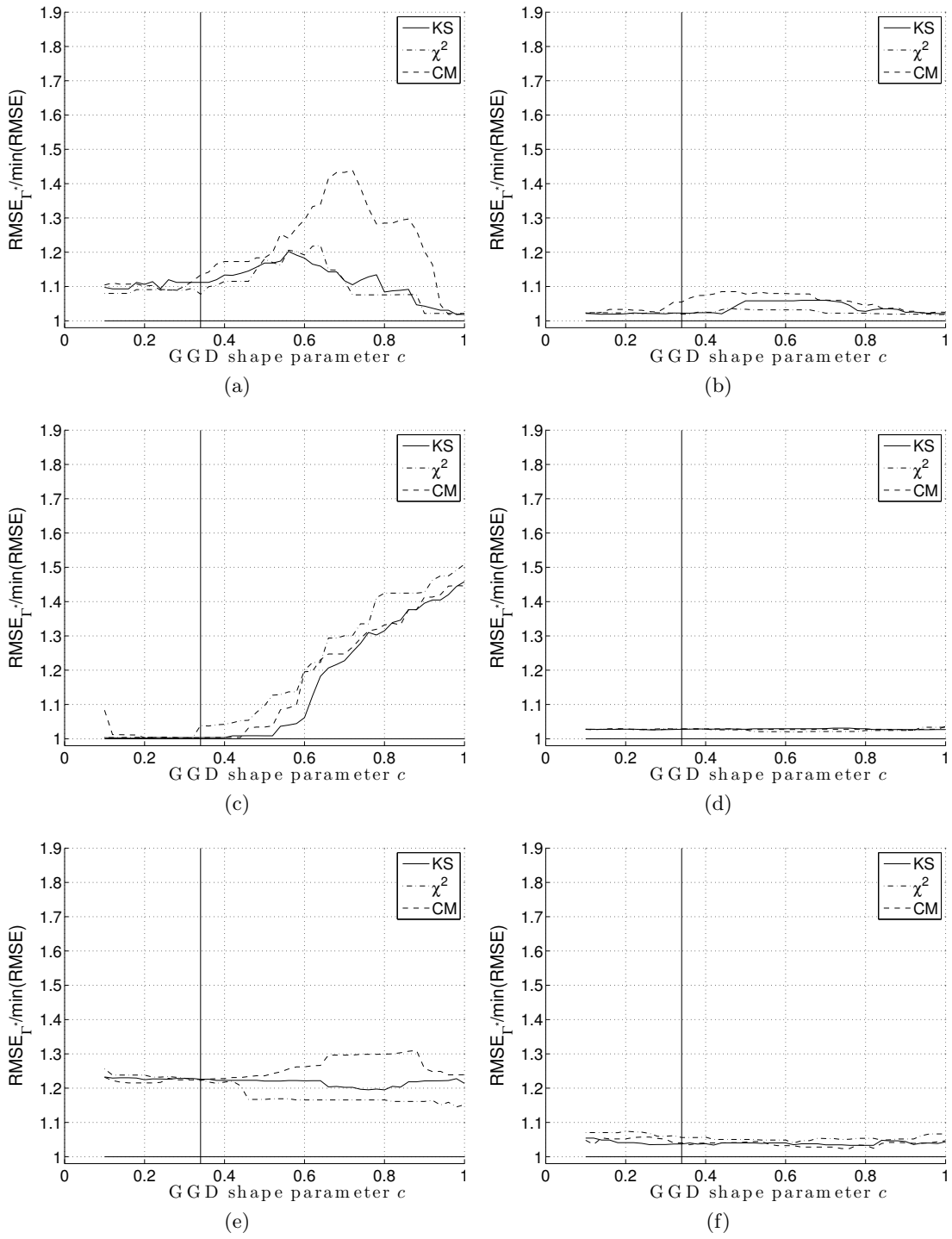


Figure C.11: Performance of the Γ selection method for 1024 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

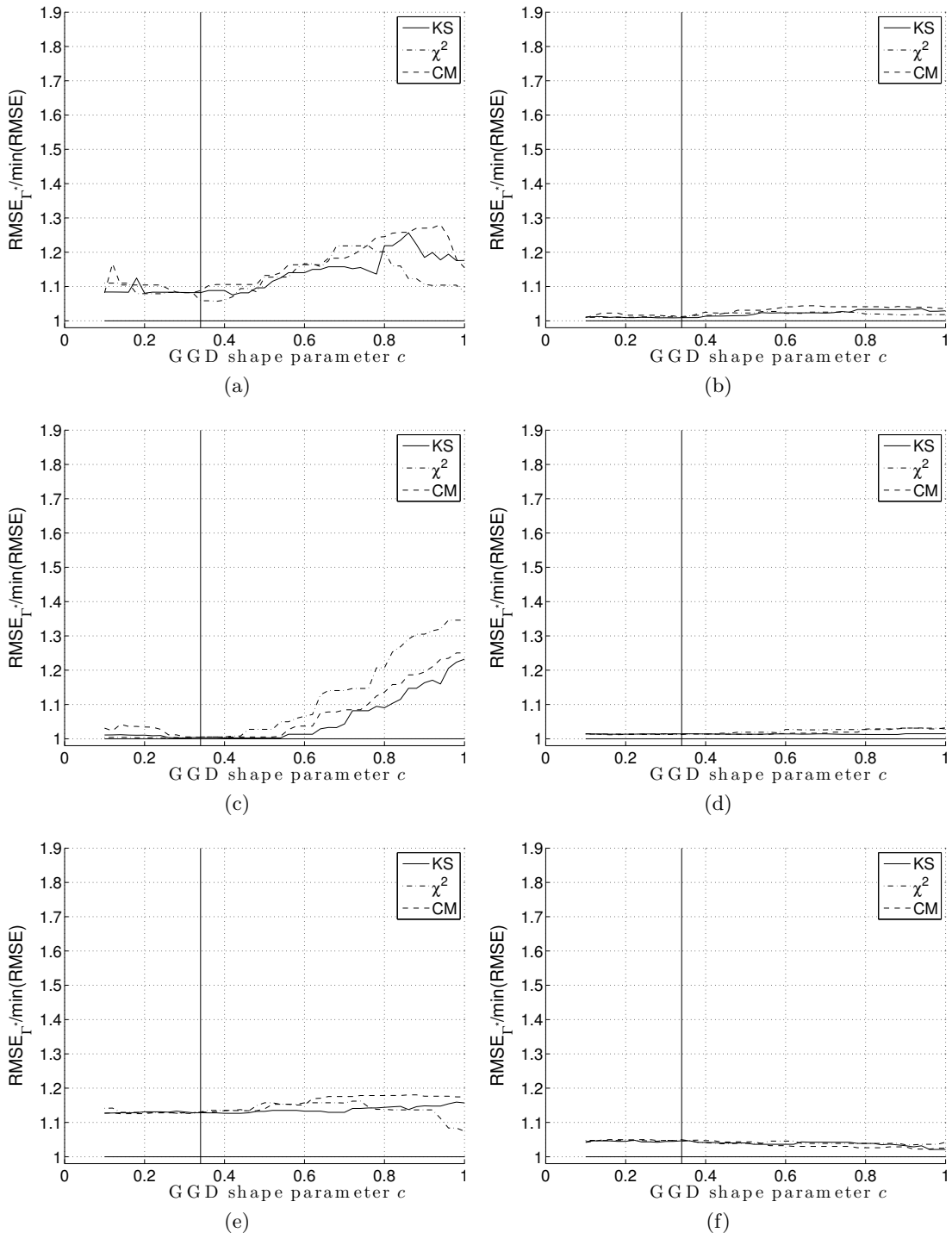


Figure C.12: Performance of the Γ selection method for 1024 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

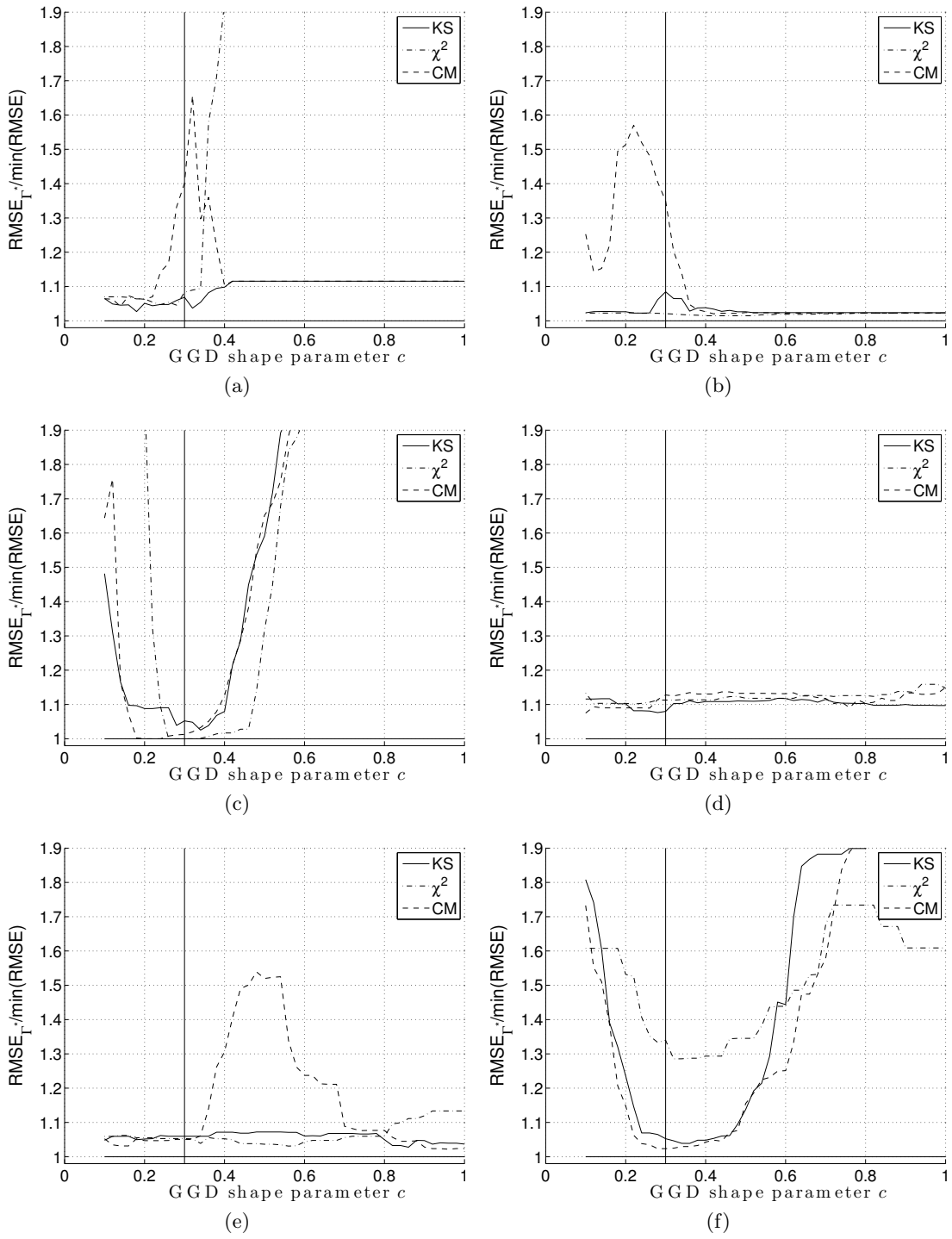


Figure C.13: Performance of the Γ selection method for 2048 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

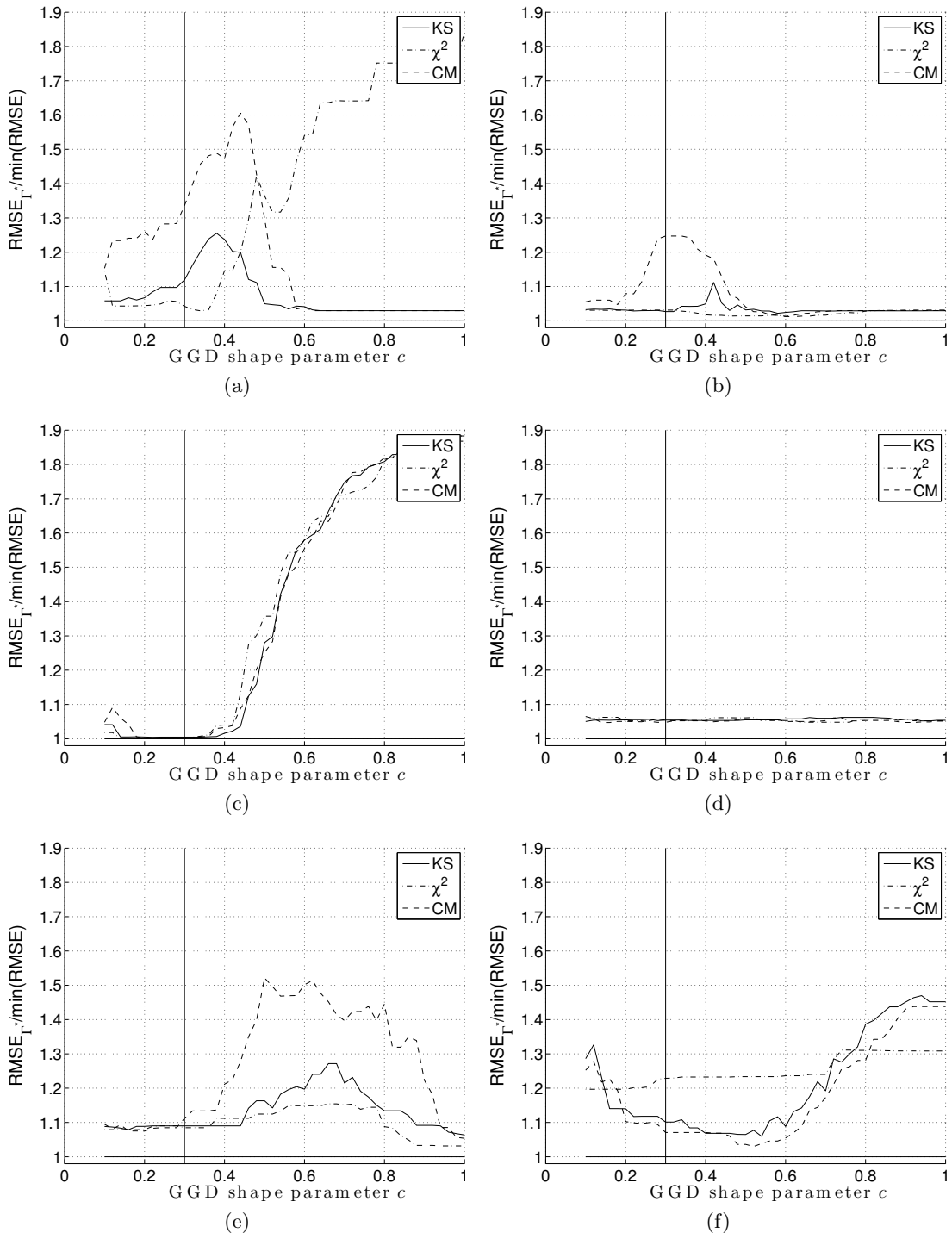


Figure C.14: Performance of the Γ selection method for 2048 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

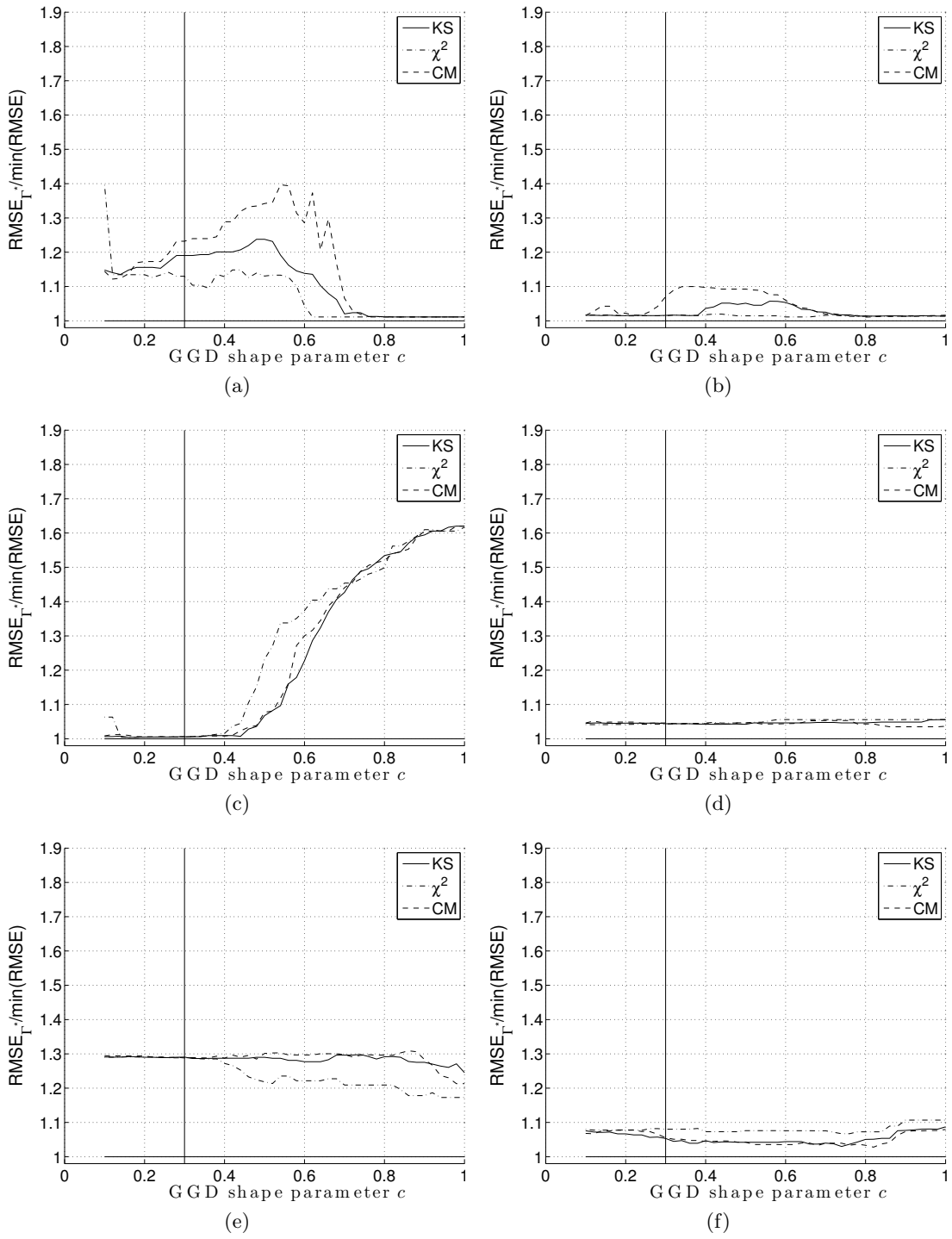


Figure C.15: Performance of the Γ selection method for 2048 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

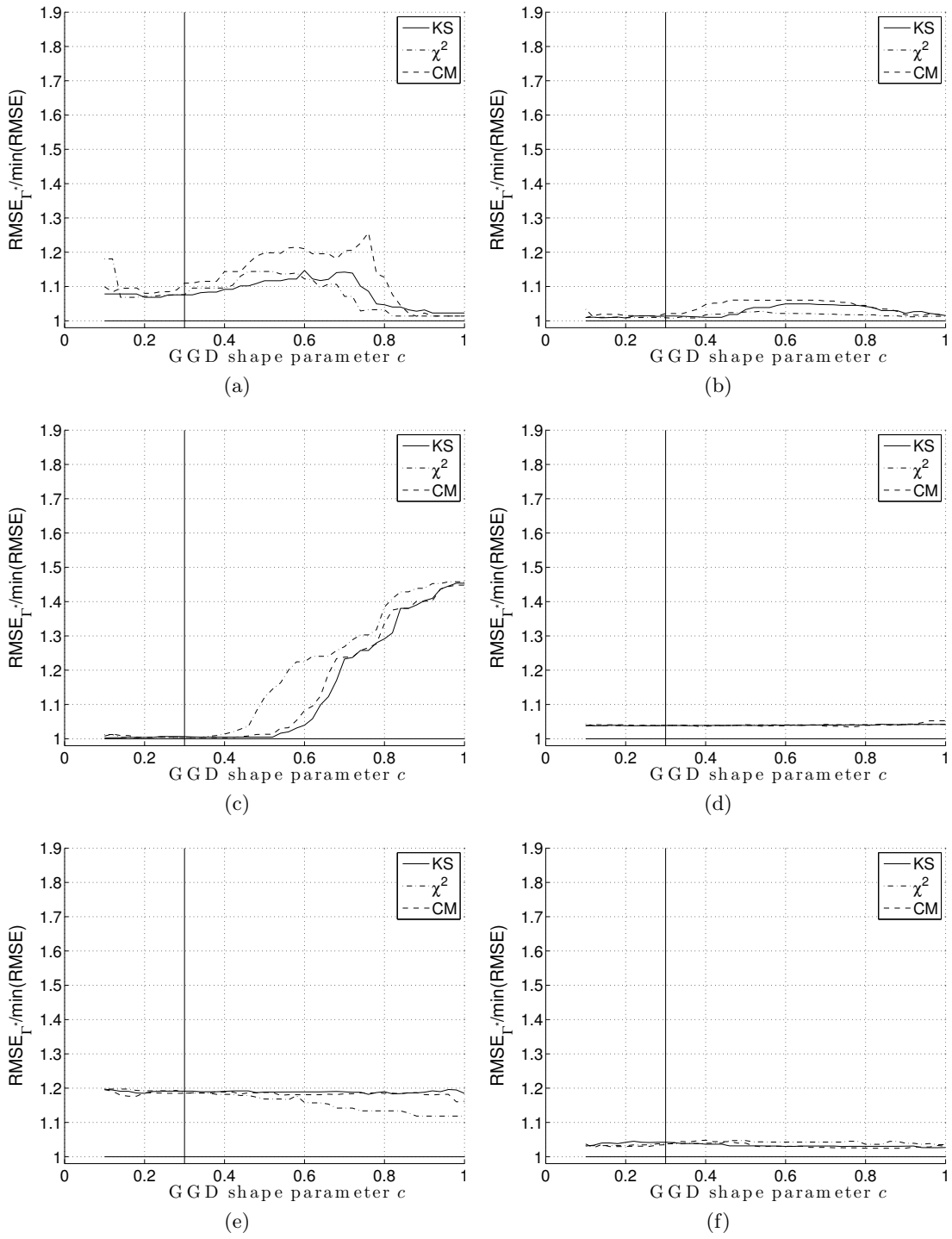


Figure C.16: Performance of the Γ selection method for 2048 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

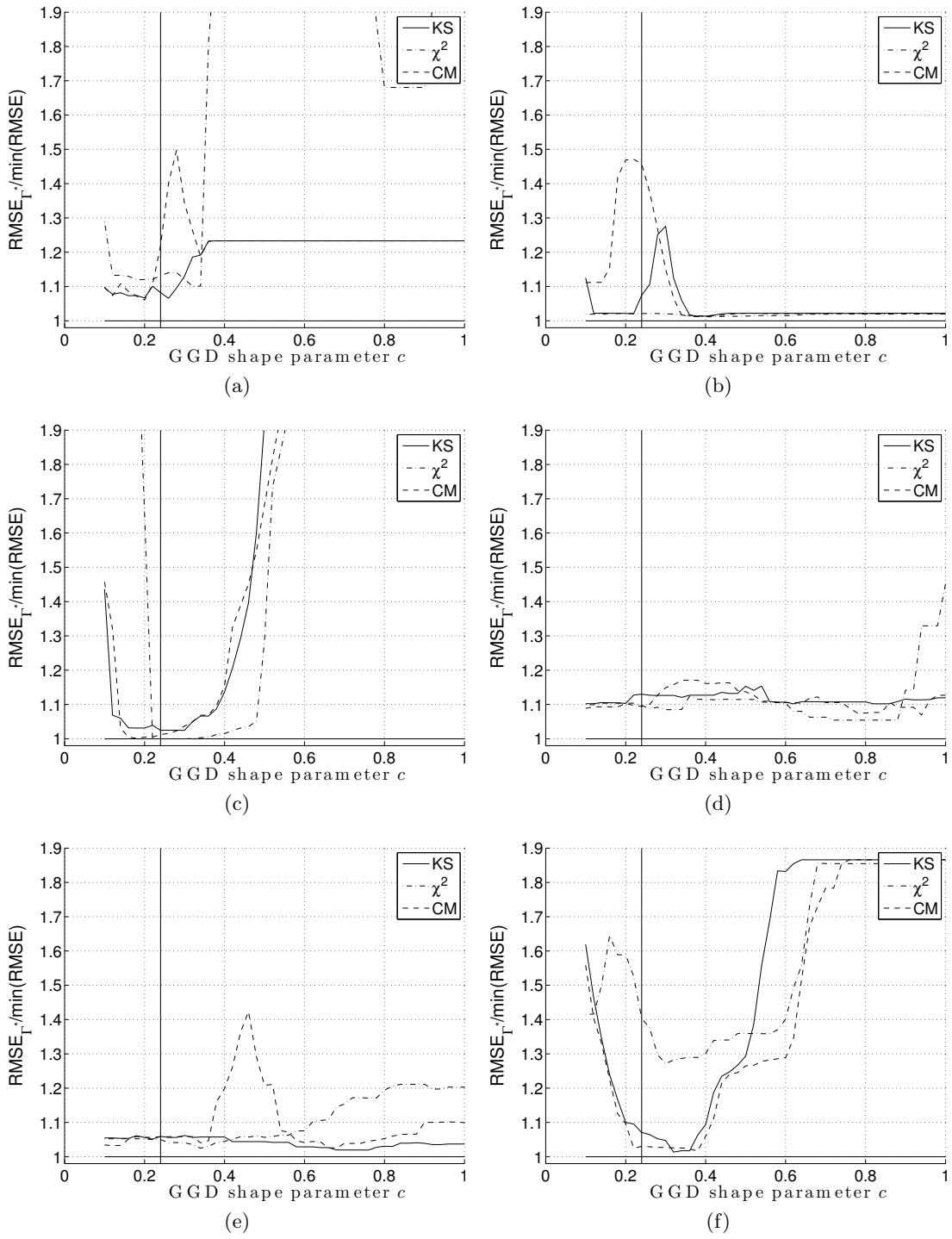


Figure C.17: Performance of the Γ selection method for 4096 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 5\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

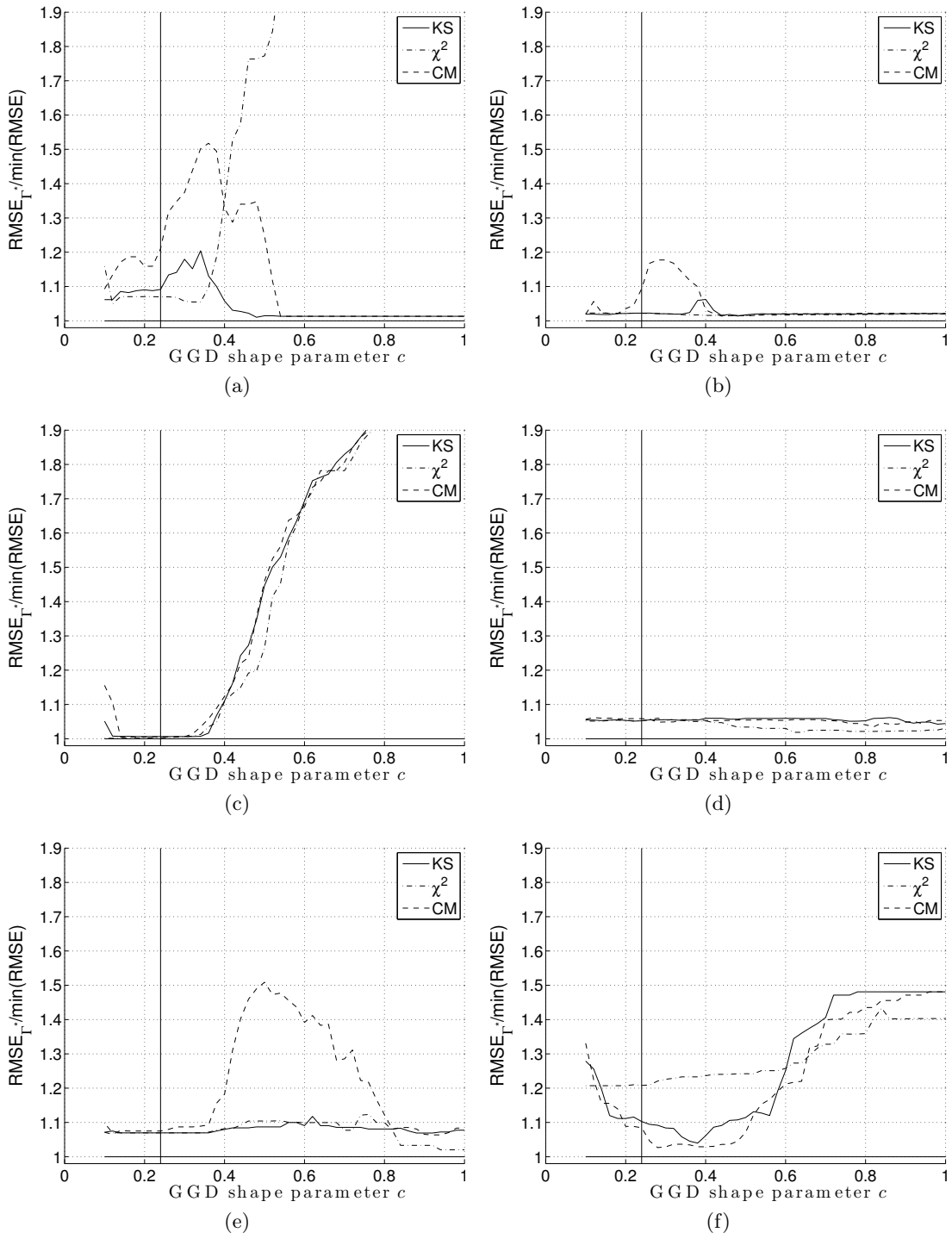


Figure C.18: Performance of the Γ selection method for 4096 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 10\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

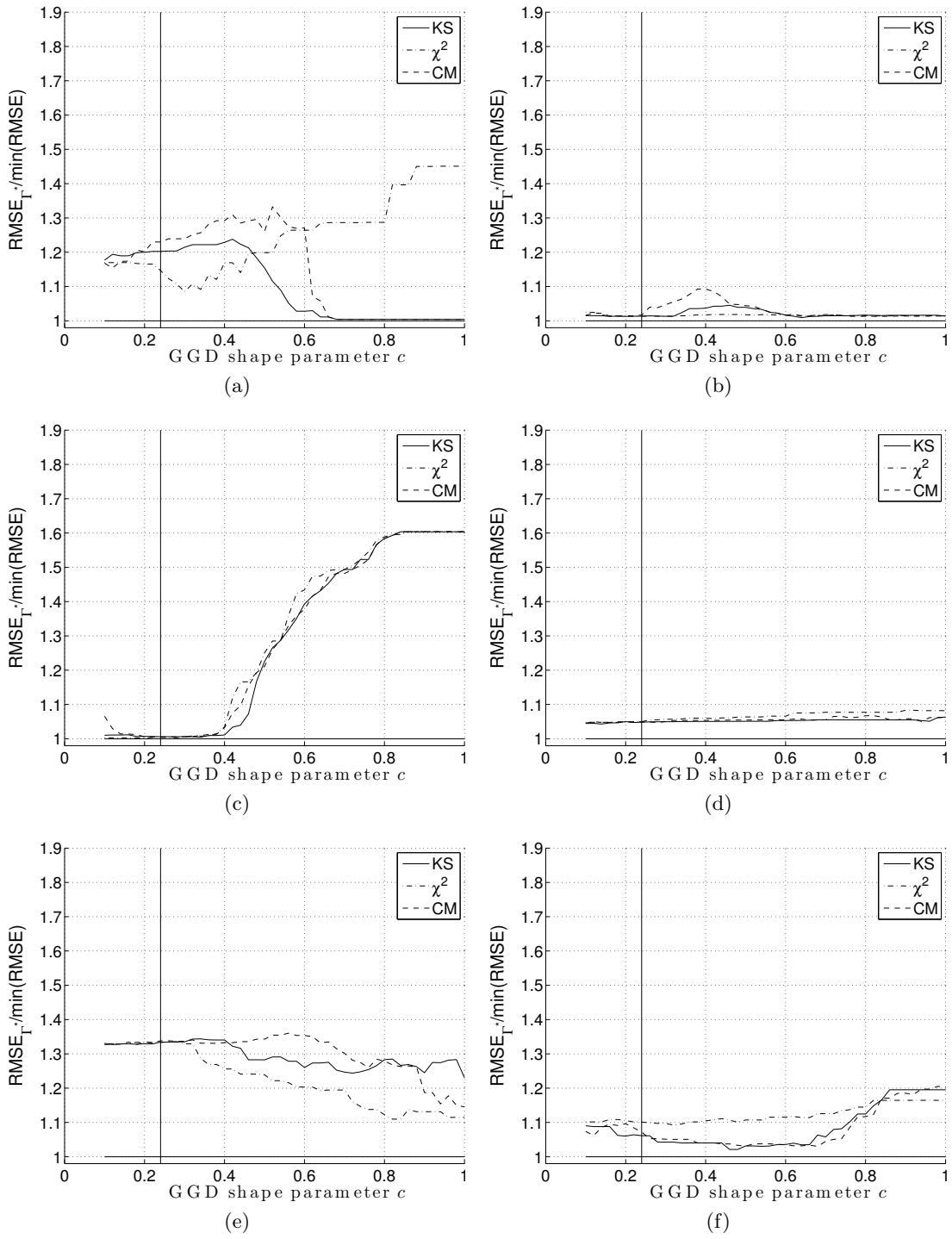


Figure C.19: Performance of the Γ selection method for 4096 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 15\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

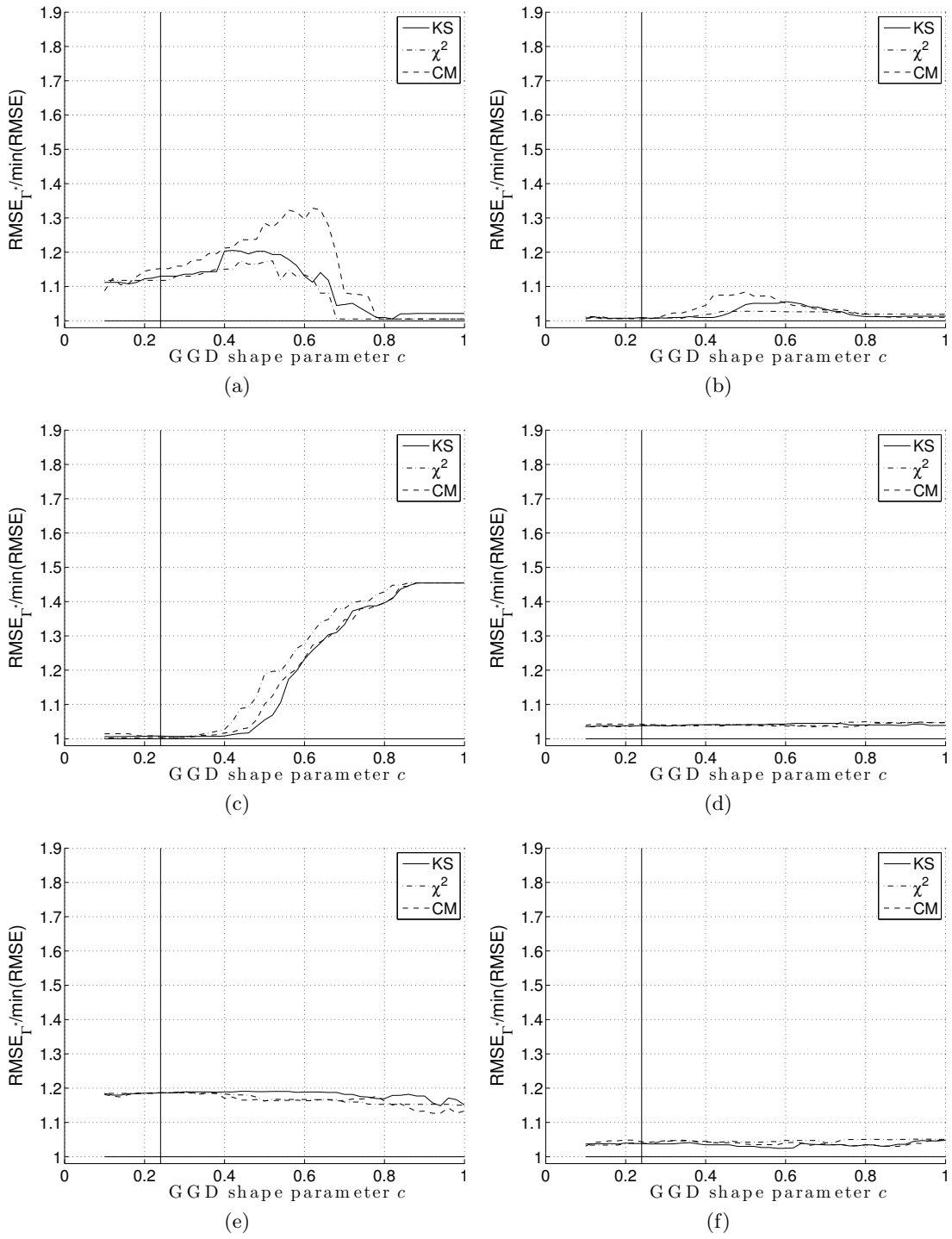


Figure C.20: Performance of the Γ selection method for 4096 samples long signals, corrupted by additive white Gaussian noise with $\sigma_\epsilon = 20\%$ of total signal magnitude. Appropriate GGD shape parameter value is denoted by the vertical line. Test signals: *Blocks* (a), *Bumps* (b), *Doppler* (c), *HeaviSine* (d), *Piece-Polynomial* (e) and *Piece-Regular* (f)

Bibliography

Abramowitz, M. and I. A. Stegun (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (ninth Dover printing, tenth GPO printing ed.). New York: Dover.

Anderson, T. W. (1962). On the distribution of the two-sample cramer-von mises criterion. *Annals of Mathematical Statistics 33*(3), 1148–1159.

Aufrichtig, R. and D. Wilson (1995). X-ray fluoroscopy spatio-temporal filtering with object detection. *IEEE Trans. Med. Imag. 14*(4), 733–746.

Berlin, L. (2001, July). Radiation-induced skin injuries and fluoroscopy. *Am. J. Roentgenol. 177*(1), 21–25.

Bismuth, V. and R. Vaillant (2008). A device enhancing and denoising algorithm for x-ray cardiac fluoroscopy. In *Proc. 19th International Conference on Pattern Recognition ICPR 2008*, pp. 1–4.

Burrus, C., R. A. Gopinath, and H. Guo (1998). *Introduction to wavelets and wavelet transforms : a primer*. Upper Saddle River, N.J: Prentice Hall.

Chakravarti, I. M., R. G. Laha, and J. Roy (1967). Handbook of methods of applied statistics. *I*.

Chan, T. F. and H. M. Zhou (1999). Adaptive ENO-wavelet transforms for discontinuous functions. Technical Report 99-21.

Claypoole, R., R. Baraniuk, and R. Nowak (1998). Adaptive wavelet transforms via lifting. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Volume 3, pp. 1513–1516 vol.3.

Claypoole, R., G. Davis, W. Sweldens, and R. Baraniuk (1997). Nonlinear wavelet transforms for image coding. In *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, Volume 1, pp. 662–667 vol.1.

Claypoole, R., G. Davis, W. Sweldens, and R. Baraniuk (2003). Nonlinear wavelet transforms for image coding via lifting. *IEEE Trans. Image Process. 12*(12), 1449–1459.

Coifman, R. R. and D. L. Donoho (1996). Translation-invariant de-noising.

Daubechies, I. (1992, December). *Ten Lectures on Wavelets (C B M S - N S F Regional Conference Series in Applied Mathematics)*. Soc for Industrial & Applied Math.

Daubechies, I. and W. Sweldens (1998). Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl. 4*(3), 245–267.

Donoho, D. L. and I. M. Johnstone (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika 81*(3), 425–455.

Donoho, D. L. and I. M. Johnstone (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association 90*(432), 1200–1224.

Donoho, D. L. and I. M. Johnstone (1998). Minimax estimation via wavelet shrinkage. *Annals of Statistics 26*(3), 879–921.

Donoho, D. L., I. M. Johnstone, G. Kerkyacharian, and D. Picard (1995). Wavelet shrinkage: Asymptopia? *J. R. Statist. Soc. B. 57*(2), 301–337.

Foi, A., V. Katkovnik, and K. Egiazarian (2007). Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images. *IEEE Trans. Image Process. 16*(5), 1395–1411.

Frazier, T., J. Richardson, V. Fabre, and J. Callen (2007). Fluoroscopy-induced chronic radiation skin injury: a disease perhaps often overlooked. *Archives of Dermatology 143*(5), 637–640.

Katkovnik, V. (1999). A new method for varying adaptive bandwidth selection. *IEEE Transactions on Signal Processing 47*(9), 2567–2571.

Katkovnik, V., K. Egiazarian, and J. Astola (2002). Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule. *J. Math. Imaging Vis. 16*(3), 223–235.

Katkovnik, V., K. Egiazarian, and J. Astola (2003). *Adaptive varying scale methods in image processing*. Tampere lntemational Center for Signal Processing, TlCSP Series,N 19, Tampere, TTY, Monistamo.

Koenig, T., F. Mettler, and L. Wagner (2001). Skin injuries from fluoroscopically guided procedures : Part 2, review of 73 cases and recommendations for minimizing dose delivered to patient. *American journal of roentgenology 177*(1), 13–20.

Kovacevic, J. and W. Sweldens (2000). Wavelet families of increasing order in arbitrary dimensions. *IEEE Transactions on Image Processing 9*(3), 480–496.

Lam, E. (2004). Statistical modelling of the wavelet coefficients with different bases and decomposition levels. *IEE Proceedings -Vision, Image and Signal Processing 151*(3), 203–206.

Mallat, S. (1988). Multiresolution representations and wavelets.

Mallat, S. (1989a). Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics, Speech, and Signal Processing 37*, 2091–2110.

Mallat, S. (1989b). Multiresolution approximation and wavelet orthogonal bases of l2(r). *Transactions of the American Mathematical Society 315*.

Mallat, S. G. (1989c). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell. 11*(7), 674–693.

Manjeshwar, R. and P. Dhawale (2005). Quantitative fluoroscopic dose saving in cardiovascular imaging with a novel motion discriminating temporal filter. In *Proc. Computers in Cardiology*, pp. 227–230.

Piella, G. and H. Heijmans (2002). Adaptive lifting schemes with perfect reconstruction.

Piella, G., B. Pesquet-Popescu, and H. Heijmans (2002). Adaptive update lifting with a decision rule based on derivative filters. *IEEE Signal Processing Letters 9*(10), 329–332.

Schoonenberg, G., M. Schrijver, Q. Duan, R. Kemkers, and A. Laine (2005). Adaptive spatial-temporal filtering applied to x-ray fluoroscopy angiography. Volume 5744, pp. 870–878. SPIE.

Sersic, D. (2000a). Adaptation criterions for wavelet filter banks with variable zero moments. In *Proc. 10th Mediterranean Electrotechnical Conference MELECON 2000*, Volume 2, pp. 603–606 vol.2.

Sersic, D. (2000b). A realization of wavelet filter bank with adaptive filter parameters. In *Proc. of the European Signal Processing Conference*, Volume 3, pp. 1733–1736.

Sersic, D. (2000c). Wavelet filter banks with adaptive number of zero moments. In *Proc. 5th International Conference on Signal Processing WCCC-ICSP 2000*, Volume 1, pp. 325–328 vol.1.

Sersic, D. and M. Vrankic (2002). Adaptation of a 2-d nonseparable wavelet filter bank with variable number of zero moments. In *Proceedings of the Second IAESTED International Conference on Visualization, Imaging and Image Processing*, pp. 257–260.

Snedecor, G. W. and W. G. Cochran (1967). *Statistical methods*. Iowa State University Press.

Stankovic, L. (2004). Performance analysis of the adaptive algorithm for bias-to-variance tradeoff. *IEEE Transactions on Signal Processing 52*(5), 1228–1234.

Stephens, M. A. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters. *Annals of Statistics 4*(2), 357–369.

Strang, G. and T. Nguyen (1996). *Wavelets and Filter Banks*. Wellesley, MA, USA: Wellesley-Cambridge Press.

Sweldens, W. (1996). The lifting scheme: A custom-design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal. 3*(2), 186–200.

Sweldens, W. (1997). The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal. 29*(2), 511–546.

Sweldens, W. and P. Schroder (1997). Building your own wavelets at home.

Thas, O. and J. P. Ottoy (2002). Goodness-of-fit tests based on sample space partitions: a unifying overview. *Journal of Applied Mathematics and Decision Sciences 6*(4), 203–212.

Thode, H. (2002). *Testing for Normality*. New York: Marcel Dekker.

Tomic, M. (2008a). Improving denoising performance with adaptive wavelet transforms. In *Proceedings of the 19th DAAAM International Symposium*.

Tomic, M. (2008b). Wavelet transforms with application in signal denoising. In *Proceedings of the 19th DAAAM International Symposium*.

Tomic, M., D. Sersic, and M. Vrankic (2008, May 22). Edge-preserving adaptive wavelet denoising using ici rule. *Electronics Letters 44*(11), 698–699.

Vetterli, M. and J. Kovacevic (1995). *Wavelets and subband coding*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.

Wang, Y., H. Li, J. Xuan, S.-C. Lo, and S. Mun (1997). Modeling of wavelet coefficients in medical image compression. In *Proc. International Conference on Image Processing*, Volume 1, pp. 644–647 vol.1.

Wu, Y., Q. Pan, H. Zhang, and S. Zhang (2004). Adaptive denoising based on lifting scheme. In *Proc. 7th International Conference on Signal Processing ICSP '04*, Volume 1, pp. 352–355 vol.1.

Abstract

We propose an adaptive lifting scheme with a goal of improving the wavelet transform performance about edges in a signal. The adaptive algorithm is based on the statistical method of intersection of confidence intervals (ICI) rule. It is used on a point-by-point basis, and on each scale. As a final result, longer and smoother wavelets are used in smooth signal regions, while shorter wavelets are used in higher frequency regions. The approach allows for efficient reconstruction of edges or, in general, higher signal frequencies. Additionally, a method for automated sensitivity parameter Γ value selection was proposed, as its proper value is a prerequisite for high transform efficiency. The adaptive algorithm was also tested in a real-world application of fluoroscopic image sequences denoising, in which edge preservation is an essential requirement. The proposed adaptive edge preserving lifting scheme and accompanying parameter Γ selection method were shown to represent a well performing model of the second generation wavelets, i.e., wavelets which inherit all the benefits and good properties of the classical wavelet transforms, while in the same time introducing additional advantages and features.

Keywords: wavelet transform, adaptive lifting scheme, second generation wavelets, local polynomial approximation, intersection of confidence intervals, signal denoising, fluoroscopic imaging

Sažetak

Predložena je adaptivna shema podizanja čiji je cilj poboljšavanje učinkovitosti valićne transformacije u okolini rubova unutar signala. Adaptivni je algoritam baziran na stastističkoj metodi pravilo presjeka intervala pouzdanosti (Intersection of Confidence Intervals – ICI). Metoda se koristi na svakom nivou razlaganja i u svakoj točki signala. Kao konačni rezultat, dulji i glatki valići koriste se za glatke dijelove signala, dok se kraći valići koriste u dijelovima signala u kojima su prisutne više frekvencije. Takav pristup omogućava učinkovitu rekonstrukciju rubova unutar signala odnosno, općenito, visokih frekvencija. Dodatno, predložena je i metoda za automatsko određivanje vrijednosti parametra osjetljivosti transformacije, Γ, čiji je dobar odabir preduvjet za njenu visoku učinkovitost. Adaptivni je algoritam testiran i na primjeni uklanjanja šuma iz stvarnih fluoroskopskih nizova slika, pri čemu je očuvanje rubova unutar slike ključan zahtjev. Predložena adaptivna shema podizanja sa svojstvom očuvanja rubova unutar signala, te pripadajuća metoda odabira Γ parametra, učinkovita su realizacija valića druge generacije, odnosno valića koji zadržavaju sve prednosti i dobre značajke klasičnih valića, ali u isto vrijeme donose i nove mogućnosti te poboljšanja učinkovitosti.

Ključne riječi: valićna transformacija, adaptivna shema podizanja, valići druge generacije, lokalna aproksimacija polinomima, presjek intervala pouzdanosti, uklanjanje šuma iz signala, fluoroskopske snimke

Biography

Mladen Tomic was born in Karlovac in 1977, where he attended both primary and secondary school, and graduated in 1995 from Gymnasium Karlovac. That autumn, he entered Faculty of Electrical Engineering and Computing at University of Zagreb, Croatia. In 1998 he was presented the *Josip Loncar* best students award and in 2000 he received a diploma in electrical engineering. The same year, Tomic took position of the research assistant at the Department of electrical engineering, Faculty of engineering at University of Rijeka, Croatia, where he still works today. His teaching activities are mainly focused on computing and computer science courses, such as: Programming, Computer System Architecture and Applied Computing.

In 2004 Tomic received the Master of Science degree for the thesis entitled: "MPEG-4 Video Decoder With Adaptive Application of Postprocessing Filters". The same year he finished the Cisco Networking Academy Instructor Curriculum and joined the local academy at the Faculty of Engineering as an instructor for the CCNA, and later, CCNAE curricula. His main research interests are digital signal processing, wavelet transforms, multimedia and computer networks.

Publications

Tomic, M. (2008a). Improving denoising performance with adaptive wavelet transforms. In *Proceedings of the 19th DAAAM International Symposium.*

Tomic, M. (2008b). Wavelet transforms with application in signal denoising. In *Proceedings of the 19th DAAAM International Symposium.*

Tomic, M., D. Sersic, and M. Vrankic (2008, May 22). Edge-preserving adaptive wavelet denoising using ici rule. *Electronics Letters 44*(11), 698–699.

Tomic, M. and N. Stojkovic (2007). Framework for adaptive postprocessing of video signals. *Automatika 48*, 21–25.

Tomic, M., N. Stojkovic, and M. Kovac (2003). Implementation and analysis of mpeg-4 dynamic resolution conversion. In *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis ISPA 2003*, pp. 317–322.

Životopis

Mladen Tomić rođen je u Karlovcu, 1977. godine, gdje je i pohađao osnovnu i srednju školu. Maturirao je 1995. godine u Gimnaziji Karlovac. Iste godine upisuje Fakultet elektrotehnike i računarstva, Sveučilišta u Zagrebu. Godine 1998. za uspjeh na 3. godini studija dobiva priznanje "Josip Lončar". Diplomirao je 2000. godine te stekao zvanje diplomiranog inženjera elektrotehnike. Iste godine zapošljava se kao znanstveni novak na Tehničkom fakultetu Sveučilišta u Rijeci, gdje i danas radi. Nastavne aktivnosti su mu uglavnom usmjerene na kolegije iz područja računarstva i računarske znanosti, poput Programiranja, Arhitekture računalnih sustava i Primjene računala.

Godine 2004. uspješno brani magistarski rad pod naslovom: "Video dekoder prema MPEG-4 normi s adaptivnom primjenom filtra za postprocesiranje", te stiče zvanje magistra znanosti računarstva. Iste godine završava i instruktorski program Cisco akademije mrežnih tehnologija te se pridružuje lokalnoj akademiji na Tehničkom fakultetu kao instruktor CCNA, odnosno kasnije, CCNAE programa. Osnovna su mu područja interesa digitalna obrada signala, valićne transformacije, multimedija te računalne mreže.

Objavljeni radovi

Tomic, M. (2008a). Improving denoising performance with adaptive wavelet transforms. In *Proceedings of the 19th DAAAM International Symposium*.

Tomic, M. (2008b). Wavelet transforms with application in signal denoising. In *Proceedings of the 19th DAAAM International Symposium*.

Tomic, M., D. Sersic, and M. Vrankic (2008, May 22). Edge-preserving adaptive wavelet denoising using ici rule. *Electronics Letters 44*(11), 698–699.

Tomic, M. and N. Stojkovic (2007). Framework for adaptive postprocessing of video signals. *Automatika 48*, 21–25.

Tomic, M., N. Stojkovic, and M. Kovac (2003). Implementation and analysis of mpeg-4

dynamic resolution conversion. In *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis ISPA 2003*, pp. 317–322.