

Article

# Properties of the SURE Estimates When Using Continuous Thresholding Functions for Wavelet Shrinkage

Alexey Kudryavtsev <sup>1,2,\*</sup>  and Oleg Shestakov <sup>1,2,3,\*</sup> 

<sup>1</sup> Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, Moscow 119991, Russia

<sup>2</sup> Moscow Center for Fundamental and Applied Mathematics, Moscow 119991, Russia

<sup>3</sup> Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow 119333, Russia

\* Correspondence: aakudryavtsev@cs.msu.ru (A.K.); oshestakov@cs.msu.ru (O.S.)

**Abstract:** Wavelet analysis algorithms in combination with thresholding procedures are widely used in nonparametric regression problems when estimating a signal function from noisy data. The advantages of these methods lie in their computational efficiency and the ability to adapt to the local features of the estimated function. It is usually assumed that the signal function belongs to some special class. For example, it can be piecewise continuous or piecewise differentiable and have a compact support. These assumptions, as a rule, allow the signal function to be economically represented on some specially selected basis in such a way that the useful signal is concentrated in a relatively small number of large absolute value expansion coefficients. Then, thresholding is performed to remove the noise coefficients. Typically, the noise distribution is assumed to be additive and Gaussian. This model is well studied in the literature, and various types of thresholding and parameter selection strategies adapted for specific applications have been proposed. The risk analysis of thresholding methods is an important practical task, since it makes it possible to assess the quality of both the methods themselves and the equipment used for processing. Most of the studies in this area investigate the asymptotic order of the theoretical risk. In practical situations, the theoretical risk cannot be calculated because it depends explicitly on the unobserved, noise-free signal. However, a statistical risk estimate constructed on the basis of the observed data can also be used to assess the quality of noise reduction methods. In this paper, a model of a signal contaminated with additive Gaussian noise is considered, and the general formulation of the thresholding problem with threshold functions belonging to a special class is discussed. Lower bounds are obtained for the threshold values that minimize the unbiased risk estimate. Conditions are also given under which this risk estimate is asymptotically normal and strongly consistent. The results of these studies can provide the basis for further research in the field of constructing confidence intervals and obtaining estimates of the convergence rate, which, in turn, will make it possible to obtain specific values of errors in signal processing for a wide range of thresholding methods.

**Keywords:** wavelets; threshold processing; adaptive threshold; unbiased risk estimate

**MSC:** 62G20



**Citation:** Kudryavtsev, A.; Shestakov, O. Properties of the SURE Estimates When Using Continuous Thresholding Functions for Wavelet Shrinkage. *Mathematics* **2024**, *12*, 3646. <https://doi.org/10.3390/math12233646>

Academic Editor: Jüri Majak

Received: 24 October 2024

Revised: 18 November 2024

Accepted: 19 November 2024

Published: 21 November 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Wavelet methods are widely used for various signal and image processing tasks. They are applied for the frequency–time analysis of signal spectra, the identification of singular points and data compression. Moreover, during the process of receiving, processing and transmitting, signals are often contaminated with noise due to system errors, environmental interference and the internal noise of receivers. To solve this problem, various noise reduction methods have been proposed, such as linear and median filtering. Recently, wavelet decomposition has become a popular tool in solving noise reduction problems.

Compared with traditional linear and median filters, wavelet analysis methods offer advantages such as computational speed and adaptive spectral analysis capabilities: a flexible time–frequency window is automatically narrowed when considering high-frequency phenomena and expanded when studying low-frequency oscillations. This makes wavelets an effective tool for the analysis of non-stationary signals with varying regularity in their domain. Telecommunication channels and computer networks may contain additive Gaussian noise caused by thermal sources, solar radiation and other natural causes. Sensor noise, electronic component noise and poor lighting are also sources of Gaussian noise in digital images. The use of standard filtering methods may lead to the excessive blurring of small-scale image details. The use of wavelets in this situation also offers certain advantages, since applying the wavelet transform retains the Gaussian noise structure, and the useful signal is economically represented as a small number of wavelet coefficients. Thus, in the space of wavelet coefficients, it is possible to effectively separate the noise from the useful signal.

Many methods based on wavelet theory have been proposed [1]. Among these methods, threshold processing is the most widely used due to its simplicity and efficiency. Donoho and Johnstone [2] suggested that only a small number of wavelet coefficients are needed to restore the signal, and they proposed the classical method of hard threshold processing. When using the hard threshold processing function, coefficients with absolute values lower than the threshold are removed, while large coefficients remain unchanged. This function is discontinuous, and additional artifacts, such as an analog of the Gibbs effect, may appear in the restored signal. To make the threshold function continuous, Donoho proposed a soft threshold processing function that reduces large coefficients by the threshold value. However, this introduces an additional bias that can lead to distortions in the reconstructed signal (for more information, see Ref. [3]). Many authors have proposed various versions of thresholding functions that improve the classical methods to some extent [4–15]. Such functions, like the soft thresholding function, are continuous and allow one to construct an unbiased estimate of the mean square risk using Stein’s method [16]. In this case, the threshold value is usually proposed to be calculated by solving the problem of minimizing this estimate. Due to the appearance of such a variety of threshold processing methods, Johnstone [3] proposed to consider a fairly general class of thresholding functions that satisfy some natural requirements and allow the construction of an unbiased risk estimate. The study of risk estimate properties is also an important practical task, since this estimate allows one to quantitatively evaluate the error of the noise suppression method using only the observed data. However, while the properties of theoretical risk are well studied, the properties of its estimates have received less attention so far. The purpose of this paper is to fill this gap and study the statistical properties of thresholding risk estimates. This paper proves that, under some additional requirements for the threshold function, the unbiased estimate of the mean square risk is strongly consistent and asymptotically normal. These properties serve as the basis for the use of the risk estimate as a quality criterion for the thresholding method.

## 2. Wavelet Decomposition

For a function  $f \in L^2(\mathbf{R})$ , the wavelet decomposition is a series of the form

$$f = \sum_{j,k \in \mathbf{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k},$$

where  $\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$  are shifts and dilations of the chosen wavelet function  $\psi$ . The index  $j$  in this expansion is called the scale, and the index  $k$  is called the shift. Under certain conditions, the sequence  $\{\psi_{j,k}\}$  forms an orthonormal basis. The function  $\psi$  can be chosen so that it is sufficiently smooth and has other useful properties [1].

In what follows, we will consider functions  $f$  with support on some finite interval  $[a, b]$ , uniformly Lipschitz regular with the exponent  $\gamma > 0$ . If a wavelet function is  $M$  times continuously differentiable ( $M \geq \gamma$ ), has  $M$  vanishing moments, i.e.,

$$\int_{-\infty}^{\infty} x^k \psi(x) dx = 0, \quad k = 0, \dots, M - 1,$$

and rapidly decays at infinity together with its derivatives, i.e., for all  $0 \leq k \leq M$  and any  $m \in \mathbf{N}$ , there is a constant  $C_m$  such that, for all  $x \in \mathbf{R}$ ,

$$|\psi^{(k)}(x)| \leq \frac{C_m}{1 + |x|^m},$$

then

$$|\langle f, \psi_{j,k} \rangle| \leq \frac{C_f}{2^{j(\gamma+1/2)}},$$

where  $C_f$  is some positive constant depending on  $f$  (a wavelet function with the given properties always exists—for example, a Daubechies wavelet of the corresponding order). This inequality shows that, in the space of wavelet coefficients, Lipschitz regular functions have a sparse representation.

In practice, functions are defined by their values on some grid of points. Suppose that this grid is uniform and the number of points is  $N = 2^J$  for some  $J > 0$ . In the case where the number of observations differs from a power of two, some extension of the array to the nearest power of 2 is applied [17]. The discrete wavelet transform is the multiplication of the vector of function values by the orthogonal matrix  $W$  defined by the wavelet function  $\psi$ . This results in a set of discrete wavelet coefficients  $\mu_{j,k}$ , for which, for a sufficiently large  $N$ , the relation  $\mu_{j,k} \approx \sqrt{N} \langle f, \psi_{j,k} \rangle$  is valid [1]. Thus, for  $\mu_{j,k}$ , the following inequality holds:

$$|\mu_{j,k}| \leq \frac{C_f 2^{J/2}}{2^{j(\gamma+1/2)}}. \tag{1}$$

### 3. Thresholding of Wavelet Coefficients

Real observations are usually contaminated by noise. In many cases, it is assumed that the noise is additive, white and Gaussian. Thus, we consider the following model:

$$X_i = f_i + z_i, \quad i = 1, \dots, N,$$

where  $f_i$  are values of the function  $f$ , and the random variables  $z_i$  have the distribution  $N(0, \sigma^2)$  and are independent. Since the matrix  $W$  is orthogonal, the discrete wavelet coefficients have the form

$$Y_{j,k} = \mu_{j,k} + \epsilon_{j,k}, \quad j = 0, \dots, J - 1, \quad k = 0, \dots, 2^j - 1,$$

where  $\epsilon_{j,k}$  have distribution  $N(0, \sigma^2)$  and are independent. Classical methods of signal processing assume the additivity and Gaussianity of the noise. However, there are approaches (see, for example, [18,19]) that assume non-Gaussianity. The consideration of such a problem could be the subject of a separate study.

To suppress noise, a thresholding function is applied to the wavelet coefficients, the purpose of which is to remove sufficiently small coefficients that are considered noise. The most common are the hard thresholding function

$$\rho_H(y, T) = \begin{cases} y & \text{for } |y| > T; \\ 0 & \text{for } |y| \leq T \end{cases}$$

and the soft thresholding function

$$\rho_S(y, T) = \begin{cases} y - T & \text{for } y > T; \\ y + T & \text{for } y < -T; \\ 0 & \text{for } |y| \leq T \end{cases}$$

with some threshold  $T$ . However, each of them has its own drawbacks. The function  $\rho_H$  is discontinuous, which leads to a lack of stability and the appearance of additional artifacts, and the function  $\rho_S$  leads to the appearance of an additional bias in the estimate of the function  $f$ . A number of studies have proposed some alternative thresholding functions  $\rho(y, T)$ , which, in fact, are a compromise between hard and soft thresholding [4–15]. These functions are continuous, like the function  $\rho_S(y, T)$ , but  $\rho(y, T) \rightarrow y$  as  $|y| \rightarrow \infty$ , i.e., for large absolute values of the coefficients, they resemble the hard thresholding function.

Due to the emergence of various types of thresholding functions, it makes sense to consider some general classes of such functions. Let the function  $h(y, T)$  have the following properties:

1.  $h(-y, T) = -h(y, T)$  (oddness);
2.  $0 \leq h(y, T) \leq T$  for  $y \geq 0$  (boundedness);
3.  $h(T, T) = T$  (continuity).

Then, the thresholding function is defined by the expression

$$\rho_h(y, T) = \begin{cases} y - h(y, T) & \text{for } |y| > T; \\ 0 & \text{for } |y| \leq T. \end{cases}$$

This form has, for example, the soft thresholding function, the hybrid thresholding function [5,15], the thresholding function based on the hyperbolic tangent [10], the sigmoid function [12] and some others.

Thus, the signal processing scheme is as follows. First, a discrete wavelet transform is performed using the orthogonal matrix  $W$ . Then, the threshold processing of the resulting wavelet coefficients is performed, which results in the suppression of most of the noise, since the useful signal has an economical representation in the space of wavelet coefficients, and the noise is uniformly distributed over all coefficients. Finally, the inverse wavelet transform is performed using the matrix  $W^T$  to obtain the processed signal ( $W^{-1} = W^T$ , since  $W$  is orthogonal). Moreover, the task of the economical presentation of data is solved along the way for storage or transmission over communication channels.

The error (or risk) of threshold processing is defined as follows:

$$R_J(T) = \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} E \left( \rho_h(Y_{j,k}, T) - \mu_{j,k} \right)^2. \tag{2}$$

One of the key issues is the choice of the threshold value  $T$  (note that a “reasonable” threshold value should increase with  $J$  [1]). One of the first to appear was the so-called universal threshold  $T_U = \sigma\sqrt{2 \ln 2^J}$ . The following statement demonstrates that this threshold is, in a certain sense, the maximum among the “reasonable” ones [1].

**Lemma 1.** *Let  $Z_1, \dots, Z_N$  be independent random variables with distribution  $N(0, \sigma^2)$ . Then,*

$$P \left( \max_{1 \leq i \leq N} |Z_i| > \sigma\sqrt{2 \ln N} \right) \rightarrow 0$$

as  $N \rightarrow \infty$ .

Thus, the maximum noise amplitude with a high probability does not exceed the universal threshold. In Refs. [20,21], additional considerations are also given, showing that it is reasonable not to consider  $T > T_U$ .

#### 4. Statistical Estimate of the Mean Square Risk

Another approach to calculating the threshold value is based on minimizing the statistical risk estimate (2) constructed using Stein’s method [16]. If  $h(y, T)$  is differentiable with respect to the variable  $y$ , then the following relation is valid:

$$E\left(\rho_h(Y_{j,k}, T) - \mu_{j,k}\right)^2 = \sigma^2 + E g^2(Y_{j,k}) - 2\sigma^2 E g'(Y_{j,k}),$$

where  $g(Y_{j,k}) = Y_{j,k} - \rho_h(Y_{j,k}, T)$ . That is, Stein’s method consists of the fact that, in (2), the expression containing unknown  $\mu_{j,k}$  can be replaced by an expression containing only observed data.

Thus, the following expression can be used as an unbiased risk estimate:

$$\hat{R}_J(T) = \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} F(Y_{j,k}, T), \tag{3}$$

where

$$F(y, T) = \begin{cases} y^2 - \sigma^2 & \text{for } |y| \leq T; \\ h^2(y, T) + \sigma^2 - 2\sigma^2 h'_y(y, T) & \text{for } |y| > T. \end{cases} \tag{4}$$

Let  $T_S$  denote the threshold that minimizes the estimate (3):

$$\hat{R}_J(T_S) = \min_{T \in [0, T_U]} \hat{R}_J(T).$$

This threshold mimics the theoretical “ideal” threshold  $T_{\min}$  that minimizes the theoretical risk (2):

$$R_J(T_{\min}) = \min_{T \in [0, T_U]} R_J(T).$$

**Remark 1.** Since this paper studies the order of the expressions under consideration with respect to  $J$ , we will use the notations  $C, c$  and others for some positive constants that may depend on the model parameters but do not depend on  $J$ . In this case, in one chain of transformations, different constants may be designated by the same letters, and the transformations themselves are performed under the assumption that  $J$  is sufficiently large.

The following upper bound is valid for  $R_J(T_{\min})$ .

**Lemma 2.** Let  $f$  be supported on some finite interval  $[a, b]$  and be uniformly Lipschitz regular with some exponent  $\gamma > 0$ . Then,

$$R_J(T_{\min}) \leq C 2^{J/(2\gamma+1)} J^{(2\gamma+2)/(2\gamma+1)}. \tag{5}$$

**Proof.** Let the index  $j^*$  be such that

$$|\mu_{j,k}| \leq C J^{-1/2}, \quad j^* \leq j \leq J - 1.$$

In this case, due to (1),

$$j^* \approx \frac{J}{2\gamma + 1} + \frac{1}{2\gamma + 1} \log_2 J + c.$$

Let us split the expression for risk into two sums:

$$R_J(T) = \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} E(\rho_h(Y_{j,k}, T) - \mu_{j,k})^2 =$$

$$= \sum_{j=0}^{j^*-1} \sum_{k=0}^{2^j-1} E(\rho_h(Y_{j,k}, T) - \mu_{j,k})^2 + \sum_{j=j^*}^{J-1} \sum_{k=0}^{2^j-1} E(\rho_h(Y_{j,k}, T) - \mu_{j,k})^2 \equiv S_1 + S_2.$$

Let us estimate the first sum from above. Given the restrictions on  $h(y, T)$ , we obtain

$$\begin{aligned} E(\rho_h(Y_{j,k}, T) - \mu_{j,k})^2 &= E((Y_{j,k} - h(Y_{j,k}, T))\mathbb{I}(|Y_{j,k}| > T) - \mu_{j,k})^2 = \\ &= E(Y_{j,k} - h(Y_{j,k}, T) - (Y_{j,k} - h(Y_{j,k}, T))\mathbb{I}(|Y_{j,k}| \leq T) - \mu_{j,k})^2 \leq \\ &\leq 2E(Y_{j,k} - \mu_{j,k})^2 + 2E[h(Y_{j,k}, T) + (Y_{j,k} - h(Y_{j,k}, T))\mathbb{I}(|Y_{j,k}| \leq T)]^2 \leq 2\sigma^2 + CT^2. \end{aligned}$$

Hence,

$$S_1 = \sum_{j=0}^{j^*-1} \sum_{k=0}^{2^j-1} E(\rho_h(Y_{j,k}, T) - \mu_{j,k})^2 \leq C2^{j^*} T^2 \leq CT^2 2^{J/(2\gamma+1)} J^{1/(2\gamma+1)}.$$

Now, consider the second sum. Given the constraints on  $h(y, T)$  and the fact that, for each term of the second sum  $|\mu_{j,k}| \leq CJ^{-1/2}$  (and hence  $|\mu_{j,k}| \leq C/T$ , since  $T < T_U$ ), we obtain

$$\begin{aligned} E(\rho_h(Y_{j,k}, T) - \mu_{j,k})^2 &= E((Y_{j,k} - h(Y_{j,k}, T))\mathbb{I}(|Y_{j,k}| > T) - \mu_{j,k})^2 \leq \\ &\leq 2E(Y_{j,k} - h(Y_{j,k}, T))^2 \mathbb{I}(|Y_{j,k}| > T) + 2\mu_{j,k}^2 \leq \\ &\leq C \left( T \exp\left\{-\frac{(T + \mu_{j,k})^2}{2\sigma^2}\right\} + T \exp\left\{-\frac{(T - \mu_{j,k})^2}{2\sigma^2}\right\} \right) + 2\mu_{j,k}^2 \leq CT \exp\left\{-\frac{T^2}{2\sigma^2}\right\} + 2\mu_{j,k}^2. \end{aligned}$$

Taking into account (1), we have

$$S_2 = \sum_{j=j^*}^{J-1} \sum_{k=0}^{2^j-1} E(\rho_h(Y_{j,k}, T) - \mu_{j,k})^2 \leq C \left( T \exp\left\{-\frac{T^2}{2\sigma^2}\right\} 2^J + 2^{J/(2\gamma+1)} J^{-\frac{2\gamma}{2\gamma+1}} \right).$$

Thus,

$$R_J(T) = S_1 + S_2 \leq C \left( T^2 2^{J/(2\gamma+1)} J^{1/(2\gamma+1)} + T \exp\left\{-\frac{T^2}{2\sigma^2}\right\} 2^J \right).$$

In this expression, the first term increases with the growing  $T$ , and the second one decreases (while, as noted,  $T < T_U$ ). Repeating the reasoning of Ref. [22] and choosing the value

$$T^* = \sigma \sqrt{\frac{4\gamma}{2\gamma+1} \ln 2^J} - \sigma \frac{2\gamma+3}{4\gamma+2} \sqrt{\frac{2\gamma+1}{4\gamma} \frac{\ln \ln 2^J}{\sqrt{\ln 2^J}}}$$

to equalize the orders of these terms, we obtain the upper bound (5) (in this case, the value  $T^*$  represents the lower bound for the “ideal” threshold  $T_{\min}$ ). The lemma is proven.  $\square$

For the threshold value  $T_S$ , the following statement holds, showing that it cannot be too small.

**Lemma 3.** *Let  $f$  be supported on some finite interval  $[a, b]$  and be uniformly Lipschitz regular with some exponent  $\gamma > 0$ . Let, for some positive  $\theta$ ,*

$$\left| h'_y(y, T) \right| \leq CT^\theta, \quad \left| h'_T(y, T) \right| \leq CT^\theta, \quad \left| h''_{yT}(y, T) \right| \leq CT^\theta, \quad \forall y \in \mathbb{R}. \tag{6}$$

Then, for  $0 < \gamma \leq 1/2$  and arbitrary  $0 < \alpha < \frac{4\gamma}{2\gamma+1}$ ,

$$P(T_S \in [0, T_1]) \leq C2^J \exp\left\{-cJ^{-\max\{3, \theta+1\}} 2^{(1-\alpha)J}\right\}. \tag{7}$$

If  $\gamma > 1/2$ , then, for an arbitrary  $\frac{2}{3} < \alpha < \frac{2}{3} \cdot \frac{4\gamma}{2\gamma+1}$ ,

$$P(T_S \in [0, T_2]) \leq C2^J \exp\left\{-cJ^{-\max\{5/2, \theta+1/2\}}2^{(1-\alpha)J}\right\}, \tag{8}$$

where  $T_1 = \sigma\sqrt{\alpha \ln 2^J}$ ,  $T_2 = \sigma\sqrt{\frac{3\alpha}{2} \ln 2^J}$ .

**Proof.** Let  $0 < \gamma \leq 1/2$ . Then, for  $0 < \alpha < \frac{4\gamma}{2\gamma+1}$ , we obtain

$$\begin{aligned} P(T_S \in [0, T_1]) &\leq P\left(\sup_{T \in [0, T_1]} (\hat{R}_J(T_{\min}) - \hat{R}_J(T)) \geq 0\right) \leq \\ &\leq P\left(\sup_{T \in [0, T_1]} |\hat{R}_J(T_{\min}) - \hat{R}_J(T) - E(\hat{R}_J(T_{\min}) - \hat{R}_J(T))| \geq \inf_{T \in [0, T_1]} (R_J(T) - R_J(T_{\min}))\right). \end{aligned}$$

Next, since  $0 \leq h(y, T) \leq T$ ,

$$\begin{aligned} E(\rho_h(Y_{J-1,k}, T) - \mu_{J-1,k})^2 &= E((Y_{J-1,k} - h(Y_{J-1,k}, T))\mathbb{I}(|Y_{J-1,k}| > T) - \mu_{J-1,k})^2 \geq \\ &\geq E(|Y_{J-1,k}| - T)^2\mathbb{I}(|Y_{J-1,k}| > T) - 2|\mu_{J-1,k}|E|Y_{J-1,k}|\mathbb{I}(|Y_{J-1,k}| > T) + \mu_{J-1,k}^2. \end{aligned}$$

Using the inequality

$$1 - \Phi(y) > \varphi(y)\left(\frac{1}{y} - \frac{1}{y^3}\right), \quad y > 0,$$

where  $\varphi(y)$  and  $\Phi(y)$  are the density and distribution function of the standard normal law, and given that, by (1),

$$|\mu_{J-1,k}| \leq \frac{C}{2^{\gamma J}},$$

we conclude that there is a constant  $C_T > 0$  such that, for  $C_T < T \leq T_1$ ,

$$\begin{aligned} E(|Y_{J-1,k}| - T)^2\mathbb{I}(|Y_{J-1,k}| > T) - 2|\mu_{J-1,k}|E|Y_{J-1,k}|\mathbb{I}(|Y_{J-1,k}| > T) + \mu_{J-1,k}^2 &\geq \\ &\geq \frac{c}{T}e^{-\frac{T^2}{2\sigma^2}} + \mu_{J-1,k}^2. \end{aligned}$$

If  $0 < T \leq C_T$ , then, obviously, for some  $\delta > 0$ , independent of  $J$ ,

$$E(\rho_h(Y_{J-1,k}, T) - \mu_{J-1,k})^2 \geq \delta.$$

Thus,

$$\inf_{T \in [0, T_1]} R_J(T) \geq \inf_{T \in [0, T_1]} \sum_{k=0}^{2^J-1} E(\rho_h(Y_{J-1,k}, T) - \mu_{J-1,k})^2 \geq \frac{C2^J}{T_1}e^{-\frac{T_1^2}{2\sigma^2}} = \frac{C2^{\frac{2-\alpha}{2}J}}{\sqrt{\ln 2^J}}.$$

Therefore, taking into account (5),

$$\inf_{T \in [0, T_1]} (R_J(T) - R_J(T_{\min})) \geq \frac{C2^{\frac{2-\alpha}{2}J}}{\sqrt{\ln 2^J}}.$$

Denote

$$Z(T) = \hat{R}_J(T_{\min}) - \hat{R}_J(T) - E(\hat{R}_J(T_{\min}) - \hat{R}_J(T)).$$

For  $T < T'$ ,

$$|Z(T) - Z(T')| = |\hat{R}_J(T') - E\hat{R}_J(T') - \hat{R}_J(T) + E\hat{R}_J(T)|.$$

Let  $N(T, T') = \#\{(j, k) \mid T < |Y_{j,k}| \leq T'\}$ . Then, given the form of  $F(y, T)$  and the constraints (6), we obtain the estimate

$$|Z(T) - Z(T')| \leq CT'^{\max\{2,\theta\}}N(T, T') + CT'^{\max\{2,\theta+1\}}(T' - T)2^J.$$

Let us divide the interval  $[0, T_1]$  into equal parts using points

$$T^l = l\delta_J, \quad l = 0, \dots, 2^J, \quad \delta_J = \frac{T_1}{2^J}.$$

Then,

$$A_J = \left\{ \sup_{T \in [0, T_1]} |\hat{R}_J(T_{\min}) - \hat{R}_J(T) - \mathbb{E}(\hat{R}_J(T_{\min}) - \hat{R}_J(T))| \geq \frac{C2^{\frac{2-\alpha}{2}}J}{\sqrt{\ln 2^J}} \right\} \subset D_J \cup E_J,$$

where

$$D_J = \left\{ \sup_l |Z(T^l)| > \frac{C2^{\frac{2-\alpha}{2}}J}{2\sqrt{\ln 2^J}} \right\}, \quad E_J = \left\{ \sup_l \sup_{T \in [T^l, T^{l+1}]} |Z(T) - Z(T^l)| \geq \frac{C2^{\frac{2-\alpha}{2}}J}{2\sqrt{\ln 2^J}} \right\}.$$

Using Hoeffding’s inequality [23], we obtain

$$P(D_J) \leq \sum_{l=1}^{2^J} P\left( |Z(T^l)| > \frac{C2^{\frac{2-\alpha}{2}}J}{2\sqrt{\ln 2^J}} \right) \leq C2^J \exp\left\{ -c \frac{2^{(1-\alpha)J}}{\ln 2^J T_{\min}^{\max\{4,2\theta\}}} \right\}. \tag{9}$$

Next,

$$\begin{aligned} E_J &\subset \left\{ \sup_l \sup_{T \in [T^l, T^{l+1}]} (CT^{\max\{2,\theta\}}N(T^l, T) + CT^{\max\{2,\theta+1\}}(T - T^l)2^J) \geq \frac{C2^{\frac{2-\alpha}{2}}J}{2\sqrt{\ln 2^J}} \right\} \subset \\ &\subset \left\{ \sup_l C(T^{l+1})^{\max\{2,\theta\}}N(T^l, T^{l+1}) \geq \frac{C2^{\frac{2-\alpha}{2}}J}{2\sqrt{\ln 2^J}} - C(T_1)^{\max\{3,\theta+2\}} \right\}. \end{aligned}$$

Since

$$\mathbb{E}N(T^l, T^{l+1}) = \mathbb{E} \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \mathbb{I}(T^l < |Y_{j,k}| \leq T^{l+1}) \leq CT_1,$$

we have

$$\begin{aligned} &\left\{ \sup_l C(T^{l+1})^{\max\{2,\theta\}}N(T^l, T^{l+1}) \geq \frac{C2^{\frac{2-\alpha}{2}}J}{2\sqrt{\ln 2^J}} - C(T_1)^{\max\{3,\theta+2\}} \right\} \subset \\ &\subset \left\{ \sup_l \left| C(T^{l+1})^{\max\{2,\theta\}}N(T^l, T^{l+1}) - \mathbb{E}C(T^{l+1})^{\max\{2,\theta\}}N(T^l, T^{l+1}) \right| \geq \right. \\ &\quad \left. \geq \frac{C2^{\frac{2-\alpha}{2}}J}{2\sqrt{\ln 2^J}} - C(T_1)^{\max\{3,\theta+2\}} \right\}. \end{aligned}$$

Applying Hoeffding’s inequality one more time, we obtain

$$P(E_n) \leq C2^J \exp\left\{ -c \frac{2^{(1-\alpha)J}}{\ln 2^J T_1^{\max\{4,2\theta\}}} \right\}. \tag{10}$$

Combining (9) and (10), we obtain



$$P(T_S \in [0, T_1]) \leq P(A_J) \leq C2^J \exp\left\{-c \frac{2^{(1-\alpha)J}}{\ln 2^J T_1^{\max\{4, 2\theta\}}}\right\} \leq C2^J \exp\left\{-cJ^{-\max\{3, \theta+1\}} 2^{(1-\alpha)J}\right\}.$$

Now, let  $\gamma > 1/2$ . Then, for  $0 < \alpha < 1$ , the estimate of the probability  $P(T_S \in [0, T_1])$  has exactly the same form. Next, let  $\frac{2}{3} < \alpha < \frac{2}{3} \cdot \frac{4\gamma}{2\gamma+1}$  and  $\beta = \frac{3\alpha}{2}$ . Then, repeating the previous reasoning, we obtain

$$\inf_{T \in [T_1, T_2]} (R_J(T) - R_J(T_{\min})) \geq \frac{C2^{\frac{2-\beta}{2}J}}{\sqrt{\ln 2^J}},$$

and

$$P(T_S \in [T_1, T_2]) \leq P(A'_J),$$

where

$$A'_J = \left\{ \sup_{T \in [T_1, T_2]} |\hat{R}_J(T_{\min}) - \hat{R}_J(T) - E(\hat{R}_J(T_{\min}) - \hat{R}_J(T))| \geq \frac{C2^{\frac{2-\beta}{2}J}}{\sqrt{\ln 2^J}} \right\} \subset D'_J \cup E'_J,$$

$$D'_J = \left\{ \sup_l |Z(T^l)| > \frac{C2^{\frac{2-\beta}{2}J}}{2\sqrt{\ln 2^J}} \right\},$$

$$E'_J = \left\{ \sup_l \sup_{T \in [T^l, T^{l+1}]} |Z(T) - Z(T^l)| \geq \frac{C2^{\frac{2-\beta}{2}J}}{2\sqrt{\ln 2^J}} \right\},$$

$$T^l = T_1 + l\delta_J, \quad l = 0, \dots, 2^J, \quad \delta_J = \frac{T_2 - T_1}{2^J}.$$

It can also be shown that

$$D(\hat{R}_J(T_{\min}) - \hat{R}_J(T)) \leq CT^{\max\{3, 2\theta-1\}} 2^{\frac{2-\alpha}{2}J}.$$

Now, repeating the previous reasoning, to estimate the probabilities of the events  $D'_J$  and  $E'_J$ , instead of Hoeffding’s inequality, we apply Bernstein’s inequality [23] and obtain

$$P(D'_J) \leq C2^J \exp\left\{-cJ^{-\max\{5/2, \theta+1/2\}} 2^{(2-2\beta+\alpha)J/2}\right\},$$

$$P(E'_J) \leq C2^J \exp\left\{-cJ^{-\max\{5/2, \theta+1/2\}} 2^{(2-2\beta+\alpha)J/2}\right\}.$$

Hence,

$$P(T_S \in [T_1, T_2]) \leq C2^J \exp\left\{-cJ^{-\max\{5/2, \theta+1/2\}} 2^{(2-2\beta+\alpha)J/2}\right\}.$$

Combining this estimate with the estimate of the probability  $P(T_S \in [0, T_1])$  and taking into account that  $\beta = \frac{3\alpha}{2}$ , we obtain the statement of the lemma for  $\gamma > 1/2$ .

The lemma is proven.  $\square$

### 5. Asymptotic Properties of the Mean Square Risk Estimate

In this section, we present statements describing the asymptotic properties of the risk estimate. For soft thresholding, similar results were obtained in [24,25].

**Theorem 1.** *Let  $f$  be supported on some finite interval  $[a, b]$  and be uniformly Lipschitz regular with some exponent  $\gamma > 1/2$ . Let  $h(y, T)$  satisfy the conditions (6). Then,*

$$P\left(\frac{\hat{R}_J(T_S) - R_J(T_{\min})}{\sigma^2 \sqrt{2^{J+1}}} < x\right) \rightarrow \Phi(x) \text{ while } J \rightarrow \infty. \tag{11}$$

**Proof.** Denote

$$U(T) = \hat{R}_J(T) - \hat{R}_J(T_{\min}) = \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} (F(Y_{j,k}, T) - F(Y_{j,k}, T_{\min})).$$

Let us write the numerator in the expression under the probability in (11) as

$$\hat{R}_J(T_S) - R_J(T_{\min}) = U(T_S) + \hat{R}_J(T_{\min}) - R_J(T_{\min})$$

and show that

$$\frac{\hat{R}_J(T_{\min}) - R_J(T_{\min})}{\sigma^2 \sqrt{2^{J+1}}} \implies N(0, 1), \quad J \rightarrow \infty. \tag{12}$$

Let  $V_J^2 = D(\hat{R}_J(T_{\min}) - R_J(T_{\min}))$ . Given (1) and the fact that  $DY_{j,k} = 2\sigma^4 + 4\sigma^2 \mu_{j,k}^2$ , it can be shown that

$$\lim_{J \rightarrow \infty} \frac{V_J^2}{\sigma^4 2^{J+1}} = 1.$$

Moreover, the Lindeberg condition is satisfied: for any  $\epsilon > 0$ , when  $J \rightarrow \infty$ ,

$$\frac{1}{V_J^2} \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} \mathbb{E} \left[ (F(Y_{j,k}, T_{\min}) - \mathbb{E}F(Y_{j,k}, T_{\min}))^2 \mathbb{I}(|F(Y_{j,k}, T_{\min}) - \mathbb{E}F(Y_{j,k}, T_{\min})| > \epsilon V_J) \right] \rightarrow 0,$$

since, starting from some  $J$ , the indicators become equal to zero. Therefore, (12) holds. It remains to show that

$$\frac{U(T_S)}{\sqrt{2^J}} \xrightarrow{P} 0. \tag{13}$$

For  $\frac{2}{3} < \alpha < \frac{2}{3} \cdot \frac{4\gamma}{2\gamma+1}$ , we denote  $T_2 = \sigma \sqrt{\frac{3}{2} \alpha \ln 2^J}$ . Then, for an arbitrary  $\epsilon > 0$ ,

$$\begin{aligned} & \mathbb{P} \left( \frac{|U(T_S)|}{\sqrt{2^J}} > \epsilon \right) \leq \\ & \leq \mathbb{P}(T_S \leq T_2) + \mathbb{P} \left( \frac{\sup_{T \in [T_2, T_U]} |U(T) - \mathbb{E}U(T)| + \sup_{T \in [T_2, T_U]} |\mathbb{E}U(T)|}{\sqrt{2^J}} > \epsilon \right). \end{aligned}$$

By Lemma 3, the first probability tends to zero. Moreover, repeating the reasoning from Lemma 3, we can verify that

$$\mathbb{P} \left( \frac{\sup_{T \in [T_2, T_U]} |U(T) - \mathbb{E}U(T)|}{\sqrt{2^J}} > \epsilon \right) \leq C 2^J \exp \left\{ -c J^{-\max\{1, \theta/2\}} 2^{J/2} \right\}.$$

Finally, given that  $\gamma > 1/2$ , it can be shown that

$$\begin{aligned} \frac{\sup_{T \in [T_2, T_U]} |\mathbb{E}U(T)|}{\sqrt{2^J}} & \leq \sup_{T \in [T_2, T_U]} \left( C T^{2J/(2\gamma+1)} J^{1/(2\gamma+1)} + C T^{\max\{1, \theta-1\}} e^{-\frac{T^2}{2\sigma^2}} 2^J \right) \leq \\ & \leq C J^{\max\{1/2, (\theta-1)/2\}} 2^{\frac{2-3\alpha}{4} J} \rightarrow 0. \end{aligned}$$

Thus, (13) holds and the theorem is completely proven.  $\square$

**Theorem 2.** Let  $f$  be supported on some finite interval  $[a, b]$  and be uniformly Lipschitz regular with some exponent  $\gamma > 0$ . Let  $h(y, T)$  satisfy the conditions (6). Then,

$$\frac{\hat{R}_J(T_S) - R_J(T_{\min})}{2^J} \rightarrow 0 \text{ a.s. when } J \rightarrow \infty.$$

**Proof.** As in the previous theorem, we write

$$\hat{R}_J(T_S) - R_J(T_{\min}) = U(T_S) + \hat{R}_J(T_{\min}) - R_J(T_{\min}),$$

where

$$U(T) = \hat{R}_J(T) - \hat{R}_J(T_{\min}).$$

If  $0 < \gamma \leq 1/2$ , then, for  $0 < \alpha < \frac{4\gamma}{2\gamma+1}$ , we denote  $T_1 = \sigma\sqrt{\alpha \ln 2^J}$ . If  $\gamma > 1/2$ , then, for  $\frac{2}{3} < \alpha < \frac{2}{3} \cdot \frac{4\gamma}{2\gamma+1}$ , we denote  $T_2 = \sigma\sqrt{\frac{3}{2}\alpha \ln 2^J}$ . In both situations, one of the estimates (7) or (8) is valid, the right-hand side of which we denote, in any case, by  $p_J$ .

Consider the case  $0 < \gamma \leq 1/2$  (the case  $\gamma > 1/2$  is considered similarly). We have

$$\begin{aligned} & \mathbb{P}\left(\frac{|U(T_S)|}{2^J} > \epsilon\right) \leq \\ & \leq \mathbb{P}(T_S \leq T_1) + \mathbb{P}\left(\frac{\sup_{T \in [T_1, T_U]} |U(T) - \mathbb{E}U(T)| + \sup_{T \in [T_1, T_U]} |\mathbb{E}U(T)|}{2^J} > \epsilon\right). \end{aligned}$$

For the second probability, repeating the reasoning from Lemma 3, we obtain

$$\begin{aligned} q_J \equiv \mathbb{P}\left(\frac{\sup_{T \in [T_1, T_U]} |U(T) - \mathbb{E}U(T)|}{2^J} > \epsilon\right) & \leq C2^J \exp\left\{-cJ^{-\max\{1, \theta/2\}}2^J\right\}, \\ \frac{\sup_{T \in [T_1, T_U]} |\mathbb{E}U(T)|}{2^J} & \rightarrow 0. \end{aligned}$$

Finally, using Hoeffding’s inequality, we obtain

$$r_J \equiv \mathbb{P}\left(\frac{\hat{R}_J(T_{\min}) - R_J(T_{\min})}{2^J} > \epsilon\right) \leq C \exp\left\{-cJ^{-\max\{2, \theta\}}2^J\right\}.$$

Thus,

$$\sum_{J=1}^{\infty} (p_J + q_J + r_J) < \infty$$

and the statement of the theorem follows from the Borel–Cantelli lemma.  $\square$

### 6. Examples

Here, we provide some examples of thresholding functions that satisfy the conditions of Theorems 1 and 2 and have been proposed by various authors in recent years. Some of them depend on additional parameters.

- Thresholding function based on the hyperbolic tangent function [10]: for  $\lambda > 0$ ,

$$\rho_h(y, T) = \begin{cases} \operatorname{sgn}(y)(|y| - T) + T \left( \frac{e^{\lambda \operatorname{sgn}(y)(|y| - T)/T} - e^{-\lambda \operatorname{sgn}(y)(|y| - T)/T}}{e^{\lambda \operatorname{sgn}(y)(|y| - T)/T} + e^{-\lambda \operatorname{sgn}(y)(|y| - T)/T}} \right), & |y| > T; \\ 0, & |y| \leq T. \end{cases}$$

- Hybrid thresholding function [11]: for  $\lambda > 1$ ,

$$\rho_h(y, T) = \begin{cases} \operatorname{sgn}(y)(|y| - |y|^{1-\lambda}T^\lambda), & |y| > T; \\ 0, & |y| \leq T. \end{cases}$$

- Sigmoid thresholding function [12]: for  $\lambda \geq 0$ ,

$$\rho_h(y, T) = \begin{cases} \operatorname{sgn}(y)(|y| - T) + \frac{2T}{1+e^{-\lambda \operatorname{sgn}(y)(|y|-T)}} - T, & |y| > T; \\ 0, & |y| \leq T. \end{cases}$$

- Soft-hard thresholding function [13] (this function depends on two threshold values; threshold value  $T'$  is an additional parameter):

$$\rho_h(y, T) = \begin{cases} 0, & |y| \leq T; \\ \operatorname{sgn}(y)T'(|y| - T)/(T' - T), & T < |y| \leq T'; \\ y, & |y| > T'. \end{cases}$$

- The smooth clipped absolute deviation (SCAD) penalty thresholding function [14]: for  $\lambda > 2$ ,

$$\rho_h(y, T) = \begin{cases} 0, & |y| \leq T; \\ \operatorname{sgn}(y)(|y| - T), & T < |y| \leq 2T; \\ \operatorname{sgn}(y)(\lambda|y| - |y| - \lambda T)/(\lambda - 2), & 2T < |y| \leq \lambda T; \\ y, & |y| > \lambda T. \end{cases}$$

Some other examples of thresholding functions can be found in the monograph [3] and the papers cited therein.

### 7. Conclusions

Wavelet analysis and threshold processing methods have found wide application in the signal analysis of noisy data. The advantages of these methods include high computational efficiency and the ability to adapt to the local features of the estimated function. Under some additional assumptions, it is possible to economically represent the signal function in such a way that the useful signal is concentrated in a relatively small number of coefficients with sufficiently large absolute values, after which threshold processing is performed to remove noise arising both for natural reasons (for example, due to background sources) and due to imperfections in the equipment used. Typically, the noise distribution is assumed to be additive and Gaussian. This model has been well studied in the literature. Various types of threshold processing and parameter selection strategies adapted to specific practical applications have been proposed. The analysis of errors arising from the application of threshold processing methods is an important practical task, since it allows one to evaluate the quality of both the methods themselves and the equipment used for processing. Most of the efforts in this area are devoted to studying the asymptotic order of the theoretical risk, which cannot be calculated in practice because it explicitly depends on the theoretical unobservable noise-free signal. However, a statistical risk estimate based only on the observed data can also be used to assess the quality of noise reduction methods.

In this paper, a model of a signal contaminated with additive Gaussian noise is considered, and the general formulation of the thresholding problem with threshold functions belonging to a certain special class is discussed. A number of examples of known threshold functions from the considered class are given, which are used in such problems as ECG and EEG signal processing, the processing of noisy images, electromyography, noise reduction in audio signals, seismic signal analysis, etc. The specific type of thresholding function for practical use depends on the application and the amount of available data. Tuning

additional parameters of the specified thresholding methods improves their performance in some specific applications. For example, [10,11] present numerical studies demonstrating the effectiveness of hybrid thresholding and hyperbolic tangent-based thresholding in ECG signal processing.

This paper obtains lower bounds for thresholds that minimize the unbiased risk estimate. It also provides conditions under which this risk estimate is asymptotically normal and strongly consistent. The authors plan to conduct a number of further studies in this direction. In particular, we will seek to obtain results concerning the construction of confidence intervals and estimates of the convergence rate, which, in turn, will allow us to obtain specific error values in signal processing for a wide range of threshold processing methods. It is also planned to consider similar models for dependent sample components and non-Gaussian noise.

**Author Contributions:** Conceptualization, A.K. and O.S.; methodology, A.K. and O.S.; formal analysis, A.K. and O.S.; investigation, A.K. and O.S.; writing—original draft preparation, A.K. and O.S.; writing—review and editing, A.K. and O.S.; supervision, A.K. and O.S.; funding acquisition, O.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Ministry of Science and Higher Education of the Russian Federation, project no. 075-15-2024-544.

**Data Availability Statement:** Data is contained within the article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Mallat, S. *A Wavelet Tour of Signal Processing*; Academic Press: New York, NY, USA, 1999; p. 857.
- Donoho, D.; Johnstone, I.M. Ideal Spatial Adaptation via Wavelet Shrinkage. *Biometrika* **1994**, *81*, 425–455. [CrossRef]
- Iain Johnstone, Department of Statistics, Stanford University. Available online: [https://imjohnstone.su.domains/GE\\_09\\_16\\_19.pdf](https://imjohnstone.su.domains/GE_09_16_19.pdf) (accessed on 23 October 2024).
- Gao, H.-Y. Wavelet Shrinkage Denoising Using the Non-Negative Garrote. *J. Comput. Graph. Stat.* **1998**, *7*, 469–488. [CrossRef]
- Chmelka, L.; Kozumplik, J. Wavelet-Based Wiener Filter For Electrocardiogram Signal Denoising. *Comput. Cardiol.* **2005**, *32*, 771–774.
- Poornachandra, S.; Kumaravel, N.; Saravanan, T.K.; Somaskandan, R. WaveShrink Using Modified Hyper-Shrinkage Function. In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 17–18 January 2006; pp. 30–32.
- Lin, Y.; Cai, J. A New Threshold Function for Signal Denoising Based on Wavelet Transform. In Proceedings of the 2010 International Conference on Measuring Technology and Mechatronics Automation, Changsha, China, 13–14 March 2010; pp. 200–203.
- Huang, H.-C.; Lee, T.C.M. Stabilized Thresholding with Generalized Sure for Image Denoising. In Proceedings of the 2010 IEEE International Conference on Image Processing (ICIP 2010), Hong Kong, China, 26–29 September 2010; pp. 1881–1884.
- Zhao, R.-M.; Cui, H.-M. Improved Threshold Denoising Method Based on Wavelet Transform. In Proceedings of the 2015 7th International Conference on Modelling, Identification and Control (ICMIC), Sousse, Tunisia, 18–20 December 2015; Art. 7409352; pp. 1–4.
- He, C.; Xing, J.; Li, J.; Yang, Q.; Wang, R. A New Wavelet Thresholding Function Based on Hyperbolic Tangent Function. *Math. Prob. Eng.* **2015**, *2015*, 528656. [CrossRef]
- Hussain, S.A.; Gorashi, S.M. Image Denoising based on Spatial/Wavelet Filter using Hybrid Thresholding Function. *Int. J. Comput. Appl.* **2012**, *42*, 5–13.
- He, H.; Tan, Y. A novel adaptive wavelet thresholding with identical correlation shrinkage function for ECG noise removal. *Chin. J. Electron.* **2018**, *27*, 507–513. [CrossRef]
- Gao, H.-Y.; Bruce, A.G. Waveshrink with firm shrinkage. *Stat. Sin.* **1997**, *7*, 855–874.
- Fan, J.; Li, R. Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Am. Stat. Assoc.* **2001**, *96*, 1348–1360. [CrossRef]
- Priya, K.D.; Rao, G.S.; Rao, P.S. Comparative analysis of wavelet thresholding techniques with wavelet-wiener filter on ecg signal. *Procedia Comput. Sci.* **2016**, *87*, 178–183. [CrossRef]
- Stein, C. Estimation of the Mean of a Multivariate Normal Distribution. *Ann. Stat.* **1981**, *9*, 1135–1151. [CrossRef]
- Boggess, A.; Narkowich, F. *A First Course in Wavelets with Fourier Analysis*; Prentice Hall: Upper Saddle River, NJ, USA, 2001; p. 283.

18. Antoniadis, A.; Leporini, D.; Pesquet, J.-C. Wavelet thresholding for some classes of non-Gaussian noise. *Stat. Neerl.* **2002**, *56*, 434–453. [[CrossRef](#)]
19. Kudryavtsev, A.A.; Shestakov, O.V. Asymptotically optimal wavelet thresholding in models with non-gaussian noise distributions. *Dokl. Math.* **2016**, *94*, 615–619. [[CrossRef](#)]
20. Donoho, D.; Johnstone, I.M. Adapting to Unknown Smoothness via Wavelet Shrinkage. *J. Am. Stat. Assoc.* **1995**, *90*, 1200–1224. [[CrossRef](#)]
21. Marron, J.S.; Adak, S.; Johnstone, I.M.; Neumann, M.H.; Patil, P. Exact Risk Analysis of Wavelet Regression. *J. Comput. Graph. Stat.* **1998**, *7*, 278–309. [[CrossRef](#)]
22. Kudryavtsev, A.A.; Shestakov, O.V. Asymptotic Behavior of the Threshold Minimizing the Average Probability of Error in Calculation of Wavelet Coefficients. *Dokl. Math.* **2016**, *93*, 295–299. [[CrossRef](#)]
23. Bennett, G. Probability inequalities for the sum of independent random variables. *J. Am. Stat. Assoc.* **1962**, *57*, 33–45. [[CrossRef](#)]
24. Shestakov, O.V. Asymptotic normality of adaptive wavelet thresholding risk estimation. *Dokl. Math.* **2012**, *86*, 556–558. [[CrossRef](#)]
25. Shestakov, O.V. On the Strong Consistency of the Adaptive Risk Estimator for Wavelet Thresholding. *J. Math. Sci.* **2016**, *214*, 115–118. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.