Estimation of the Hurst Exponent Using Trimean Estimators on Nondecimated Wavelet Coefficients

Chen Feng and Brani Vidakovic

Abstract—Hurst exponent is an important feature summarizing the noisy high-frequency data when the inherent scaling pattern cannot be described by standard statistical models. In this paper, we study the robust estimation of Hurst exponent based on non-decimated wavelet transforms (NDWT). The robustness is achieved by applying a general trimean estimator on non-decimated wavelet coefficients of the transformed data. The general trimean estimator is derived as a weighted average of the distribution’s median and quantiles, combining the median’s emphasis on central values with the quantiles’ attention to the extremes. The properties of the proposed Hurst exponent estimators are studied both theoretically and numerically. Compared with other standard wavelet-based methods (Veitch & Abry (VA) method, Soltani, Simard, & Boichu (SSB) method, median based estimators MEDL and MEDLA), our methods reduce the variance of the estimators and increase the prediction precision in most cases. The proposed methods are applied to a data set in high frequency pupillary response behavior (PRB) with the goal to classify individuals according to a degree of their visual impairment.

Index Terms—Hurst exponent, Time series, High-frequency data, Robust estimation, Non-decimated wavelet transforms, General trimean estimator, Median, Quantiles.

I. INTRODUCTION

HIGH-FREQUENCY, time series data from various sources often possess hidden patterns that reveal the effects of underlyng functional differences, but such patterns cannot be explained by basic descriptive statistics, traditional statistical models, or global trends. For example, the high-frequency pupillary response behavior (PRB) data collected during human-computer interaction capture the changes in pupil diameter in response to simple or complex stimuli. Researchers found that there may be underlying unique patterns hidden within complex PRB data, and these patterns reveal the intrinsic individual differences in cognitive, sensory and motor functions [1]. Yet, such patterns cannot be explained by the traditional statistical summaries of the PRB data, for the magnitude of the diameter depends on ambient light, not on the inherent eye function [2]. When the intrinsic individual functional differences cannot be described by basic statistics and trends in the noisy high-frequency data, the Hurst exponent might be an optional measure of patients’ characterization.

The Hurst exponent \( H \) quantifies the long memory as well as regularity, self-similarity, and scaling in a time series. Among models having been proposed for analyzing the self-similar phenomena, arguably the most popular is the fractional Brownian motion (fBm) first described by Kolmogorov [3] and formalized by Mandelbrot and Van Ness [4]. Its importance is due to the fact that fBm is the unique Gaussian process with stationary increments that is self-similar. Recall that a stochastic process \( \{X(t), t \in \mathbb{R}^d\} \) is self-similar with Hurst exponent \( H \) if, for any \( \lambda \in \mathbb{R}^+ \), \( X(\lambda t) \overset{d}{=} \lambda^{-H} X(t) \). Here the notation \( \overset{d}{=} \) means the equality in all finite-dimensional distributions. Hurst exponent \( H \) describes the rate at which autocorrelations decrease as the lag between two realizations in a time series increases.

A value \( H \) in the range 0-0.5 indicates a zig-zagging intermittent time series with long-term switching between high and low values in adjacent pairs. A value \( H \) in the range 0.5 to 1 indicates a time series with long-term positive autocorrelations, which preserves trends on a longer time horizon and gives a time series more regular appearance.

Multiresolution analysis is one of the many methods to estimate the Hurst exponent \( H \). An overview can be found in [5], [6], [7]. If processes possess a stochastic structure, i.e., Gaussianity and stationary increments, \( H \) becomes a parameter in a well-defined statistical model and can be estimated after taking the log of the quantity \( \mathbb{E}(d^2_j) \), where \( d^2_j \)'s are the wavelet coefficients at level \( j \). In fact, several estimation methods of \( H \) based on wavelets analysis exist. Veitch and Abry [5] suggest the estimation of \( H \) by weighted least square regression using the level-wise log of \( d^2_j \) since the variance of \( \log d^2_j \) depends on \( H \) and the level \( j \). In addition, the authors correct for the bias caused by the order of taking the logarithm and the average in \( \log d^2_j \). Soltani et al [8] defined a mid-energy as \( D_{j,k} = \left(d^2_{j,k} + d^2_{j,k+N/2}\right)/2 \), and showed that the level-wise averages of \( \log d^2_j \) are asymptotically normal and more stable, which is used to estimate parameter \( H \) by regression. The estimators in Soltani et al [8] consistently outperform the estimators in Veitch and Abry [5]. Shen et al [9] shows that the method of Soltani et al [8] yields more accurate estimators since it takes the logarithm of the mid-energy and then averages. Kang and Vidakovic [2] proposed MEDL and MEDLA methods based on non-decimated wavelets to estimate \( H \). MEDL estimates \( H \) by regressing the medians of \( \log d^2_j \) on level \( j \), while MEDLA uses the level-wise medians of \( \log \left(d^2_{j,k_1} + d^2_{j,k_2}\right)/2 \) to...
estimate $H$, where $k_1$ and $k_2$ are properly selected locations at level $j$ to guarantee the independence assumption. Both MEDL and MEDLA use the median of the derived distribution instead of the mean, because the medians are more robust to potential outliers that can occur when logarithmic transform of a squared wavelet coefficient is taken and the magnitude of coefficient is close to zero. Although median is outlier-resistant, it can behave unexpectedly as a result of its non-smooth character. The fact that the median is not “universally the best outlier-resistant estimator” provides a practical motivation for examining alternatives that are intermediate in behavior between the very smooth but outlier-sensitive mean and the very outlier-insensitive but non-smooth median. In this article, the general trimean estimator is derived as a weighted average of the median and the average of the median, combining the median’s emphasis on center values with the quantiles’ attention to the tails. Tukey’s trimean estimator [10], [11] and Gastwirth estimator [12], [13], [14] turn out to be two special cases under the general framework. We will use the general trimean estimators of the level-wise derived distributions to estimate $H$. In this paper, we are concerned with the robust estimation of Hurst exponent in one-dimensional setting, however, the methodology can be readily extended to a multidimensional case.

The rest of the paper consists of five additional sections and an appendix. Section 2 introduces the general trimean estimators and discusses two special estimators following that general framework: Section 3 describes estimation of Hurst exponent using the general trimean estimators, presents distributional results on which the proposed methods are based, and derives optimal weights that minimize the variances of the estimators. Section 4 provides the simulation results and compares the performance of the proposed methods to other standardly used, wavelet-based methods. The proposed methods are illustrated using the real PRB data in Section 5. The paper is concluded with a summary and discussion in Section 6.

II. GENERAL TRIMEAN ESTIMATORS

Let $X_1, X_2, ..., X_n$ be i.i.d. continuous random variables with pdf $f(x)$ and cdf $F(x)$. Let $0 < p < 1$, and let $x_p$ denote the $p$th quantile of $F$, so that $x_p = \inf\{x | F(x) \geq p\}$. If $F$ is monotone, the $p$th quantile is simply defined as $F(x_p) = p$.

Let $Y_p = X_{\lfloor np \rfloor}$ denote a sample $p$th quantile. Here $\lfloor np \rfloor$ denotes the greatest integer that is less than or equal to $np$. The general trimean estimator is defined as a weighted average of the distribution’s median and its two quantiles $Y_p$ and $Y_{1-p}$, for $p \in (0, 1/2)$:

$$\hat{\mu} = \frac{\alpha}{2} Y_p + (1 - \alpha) \frac{Y_{1/2} + \alpha}{2} Y_{1-p}. \quad (1)$$

The weights for the two quantiles are the same for $Y_p$ and $Y_{1-p}$, and $\alpha \in [0, 1]$. This is equivalent to the weighted sum of the median and the average of $Y_p$ and $Y_{1-p}$ with weights $1 - \alpha$ and $\alpha$:

$$\hat{\mu} = (1 - \alpha) Y_{1/2} + \alpha \left( \frac{Y_p + Y_{1-p}}{2} \right).$$

This general trimean estimator turns out to be more robust than mean but smoother than the median. To derive the asymptotic distribution of this general trimean estimator, the asymptotic joint distribution of sample quantiles is shown in Lemma 1; detailed proof can be found in [15].

**Lemma 1.** Consider $r$ sample quantiles, $Y_{p_1}, Y_{p_2}, \ldots, Y_{p_r}$, where $1 \leq p_1 < p_2 < \ldots < p_r \leq n$. If for any $1 \leq i \leq r$, $\sqrt{n}(|np_i|/n - p_i) \to 0$ is satisfied, then the asymptotic joint distribution of $Y_{p_1}, Y_{p_2}, \ldots, Y_{p_r}$ is:

$$\sqrt{n}((Y_{p_1}, Y_{p_2}, \ldots, Y_{p_r}) - (\xi_{p_1}, \xi_{p_2}, \ldots, \xi_{p_r})) \approx \mathcal{MN}(0, \Sigma),$$

where

$$\Sigma = (\sigma_{ij})_{r \times r}, \quad \sigma_{ij} = \frac{p_i (1 - p_j)}{f(x_{p_i}) f(x_{p_j})}, \quad i \leq j. \quad (2)$$

From Lemma 1, the asymptotic distribution of general trimean estimator will be normal as a linear combination of the components of the asymptotic multivariate normal distribution. The general trimean estimator itself may be defined in terms of order statistics as

$$\hat{\mu} = A \cdot y,$$

where

$$A = \left[ \frac{\alpha}{2}, 1 - \alpha, \frac{\alpha}{2} \right], \quad y = [Y_p, Y_{1/2}, Y_{1-p}]^T.$$

It can be easily verified that $\sqrt{n}((|np|/n - p) \to 0$ for $p \in (0, 1/2)$. If we denote $\xi = [\xi_p, \xi_{1/2}, \xi_{1-p}]^T$ the population quantiles, the asymptotic distribution of $y$ is

$$\sqrt{n} (y - \xi) \approx \mathcal{MN}(0, \Sigma),$$

where $\Sigma = (\sigma_{ij})_{3 \times 3}$, and $\sigma_{ij}$ follows equation (2) for $p_1 = p$, $p_2 = 1/2$, and $p_3 = 1 - p$. Therefore

$$\hat{\mu} \approx \mathcal{N}(E(\hat{\mu}) : Var(\hat{\mu})), \quad \text{with the theoretical expectation and variance being}$$

$$E(\hat{\mu}) = E(A \cdot y) = A \cdot E(y) = A \cdot \xi, \quad (3)$$

and

$$Var(\hat{\mu}) = Var(A \cdot y) = A Var(y) A^T = \frac{1}{n} A \Sigma A^T. \quad (4)$$

A. Tukey’s Trimean Estimator

Tukey’s trimean estimator is a special case of the general trimean estimators, with $\alpha = 1/2$ and $p = 1/4$ in equation (1). To compute this estimator, we first sort the data in ascending order. Next, we take the values that are one-fourth of the way up this sequence (the first quartile), half way up the sequence (i.e., the median), and three-fourths of the way up the sequence (the third quartile). Given these three values, we then form the weighted average, giving the central (median) value a weight of 1/2 and the two quartiles each a weight of 1/4. If we denote the Tukey’s trimean estimator as $\hat{\mu}_T$, then

$$\hat{\mu}_T = \frac{1}{4} Y_{1/4} + \frac{1}{2} Y_{1/2} + \frac{1}{4} Y_{3/4}.$$

The Tukey’s trimean estimator may be expressed more compactly as

$$\hat{\mu}_T = A_T \cdot y_T,$$
where

\[ A_T = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 4 \end{bmatrix}, \quad \text{and} \quad y_T = \begin{bmatrix} Y_{1/4} \\ Y_{1/2} \\ Y_{3/4} \end{bmatrix}^T. \]

It can be easily verified that \( \sqrt{n}(|pn|/n-p) \rightarrow 0 \) for \( p = 1/4, 1/2, \) and \( 3/4. \) If we denote \( \xi_T = [\xi_{1/4} \xi_{1/2} \xi_{3/4}]^T, \)
the corresponding theoretical quantiles, the asymptotic distribution of \( \hat{\mu}_T \) is

\[ \hat{\mu}_T \approx_{\text{N}} (A_T \cdot \xi_T, \frac{1}{n} A_T \Sigma_T A_T^T), \]

where \( \Sigma_T = (\sigma_{ij})_{3 \times 3} \) is the covariance matrix of the asymptotic multivariate normal distribution, and \( \sigma_{ij} \) follows equation (2) with \( p_1 = 1/4, p_2 = 1/2, \) and \( p_3 = 3/4. \)

### B. Gastwirth Estimator

As the Tukey’s estimator, the Gastwirth estimator is another special case of the general trimean estimators, with \( \alpha = 0.6 \)
and \( p = 1/3 \) in equation (1). If we denote this estimator as \( \hat{\mu}_G, \) then

\[ \hat{\mu}_G = 0.3 Y_{1/3} + 0.4 Y_{1/2} + 0.3 Y_{2/3}. \]

The Gastwirth estimator can be written as

\[ \hat{\mu}_G = A_G \cdot \xi_G, \]

where

\[ A_G = [0.3 \quad 0.4 \quad 0.3], \quad \text{and} \quad y_G = [Y_{1/3} \quad Y_{1/2} \quad Y_{2/3}]^T. \]

As in Tukey’s case, if we denote \( \xi_G = [\xi_{1/3} \xi_{1/2} \xi_{2/3}]^T \)
as the theoretical quantiles, since \( \sqrt{n}(|pn|/n-p) \rightarrow 0 \) for \( p = 1/3, 1/2, \)
and \( 2/3, \) the asymptotic distribution of \( \hat{\mu}_G \) is

\[ \hat{\mu}_G \approx_{\text{N}} (A_G \cdot \xi_G, \frac{1}{n} A_G \Sigma_G A_G^T), \]

where \( \Sigma_G = (\sigma_{ij})_{3 \times 3}, \) and \( \sigma_{ij} \) follows equation (2) with \( p_1 = 1/3, \)
\( p_2 = 1/2, \) and \( p_3 = 2/3. \)

### III. METHODS

Our proposal for robust estimation of Hurst exponent \( H \) is
based on non-decimated wavelet transforms (NDWT). Details
of NDWT can be found in [16], [17], [18]. The comparison
of NDWT and standard orthogonal discrete wavelet transforms
(DWT) has been discussed by Kang and Vidakovic [2]. NDWT
has several advantages when employed for Hurst exponent
estimation: 1) Input signals and images of arbitrary size can
be processed due to the absence of decimation; 2) as a
redundant transform, the NDWT increases the accuracy of
the scaling estimation; 3) least square regression can be fitted
to estimate \( H \) instead of weighted least square regression
since the variances of the level-wise derived distributions based
on logged NDWT coefficients do not depend on level; 4) local
scaling can be assessed due to the time-invariance property.
Of course, as we will discuss later, the dependence of coefficients
in NDWT is much more pronounced than in the case of DWT.

In \( J \)-level decomposition of a fBm of size \( N, \) a NDWT
yields \( N \times (J + 1) \) wavelet coefficients, with each level \( N \)
coefficients. At each level \( j, \) we generate \( N/2 \) mid-energies
as

\[ D_{j,k} = \left( \frac{d_{j,k}^2 + d_{j,k+N/2}^2}{2} \right)^{1/2}, \quad \text{for} \quad k = 1, 2, ..., N/2. \] (5)

The distributions of the \( D_{j,k} \) and \( \log D_{j,k} \) are derived under
the assumption that \( d_{j,k} \) and \( d_{j,k+N/2} \) are independent. Then
we apply the general trimean estimators of the level-wise
derived distributions to estimate parameter \( H. \) Note that for
fixed \( j, \) the generated \( N/2 \) mid-energies \( D_{j,k} \) and \( \log D_{j,k} \)
are not independent but their autocorrelations decay exponentially,
consequently, they possess only the short memory. At each level
\( j, \) we sample every \( M \) points from \( D_{j,k} \) and \( \log D_{j,k} \)
to form \( M \) groups, and assume that the \( (N/2)/M \) points of \( D_{j,k} \)
and \( \log D_{j,k} \) within each subgroup are independent, respectively.
The general trimean estimators are then applied on each of the
\( M \) groups. Note that \( M \) must be divisible by \( N/2. \)

- **Group 1:** \( \{ D_{j,1}, D_{j,1+M}, D_{j,1+2M}, ..., D_{j,(N/2−1)M+1} \} \)
  and \( \{ \log(D_{j,1}), \log(D_{j,1+M}), ..., \log(D_{j,(N/2−1)M+1}) \} \)

- **Group 2:** \( \{ D_{j,2}, D_{j,2+M}, D_{j,2+2M}, ..., D_{j,(N/2−1)M+2} \} \)
  and \( \{ \log(D_{j,2}), \log(D_{j,2+M}), ..., \log(D_{j,(N/2−1)M+2}) \} \)

- **Group M:** \( \{ D_{j,M}, D_{j,2M}, D_{j,3M}, ..., D_{j,N/2} \} \)
  and \( \{ \log(D_{j,M}), \log(D_{j,2M}), ..., \log(D_{j,N/2}) \} \)

### A. General Trimean of the Mid-energy (GTME) Method

For the general trimean of the mid-energy (GTME) method, we
derive the relationship between the general trimean estimator
of each mid-energy \( D_{j,k} \) groups from decomposition level
\( j. \) The GTME method is described in the following theorem:

**Theorem III.1.** Let \( \hat{\mu}_{j,i} \) be the general trimean estimator based
on

\[ \{ D_{j,i}, D_{j,i+M}, D_{j,i+2M}, ..., D_{j,(N/2−1)M+i} \} \equiv D(i, j) \]

where \( D(i, j) \) for \( 1 \leq i \leq M \) and \( 1 \leq j \leq J \) is the \( i \)th
of mid-energies at level \( j \) in a \( J \)-level NDWT of a fBm of size
\( N \) with Hurst exponent \( H. \) Then, the asymptotic distribution
of \( \hat{\mu}_{j,i} \) is normal,

\[ \hat{\mu}_{j,i} \approx_{\text{N}} \left( c(\alpha, p) \lambda_j \frac{2M}{N} f(\alpha, p) \lambda_j^2 \right), \] (6)

where

\[ c(\alpha, p) = \frac{\alpha}{2} \log \left( \frac{1}{p(1-p)} \right) + (1-\alpha) \log 2, \]

\[ f(\alpha, p) = \frac{\alpha(1-2p)(\alpha-4p)}{4p(1-p)} + 1, \text{and} \]

\[ \lambda_j = \sigma \cdot 2^{-\left(2H+1\right)j}. \]

In the previous, \( \sigma \) is the standard deviation of wavelet
coefficients from level 0. The Hurst exponent can be estimated
as

\[ H = \frac{1}{M} \sum_{i=1}^{M} \hat{H}_i = \frac{1}{M} \sum_{i=1}^{M} \left( -\hat{\lambda}_i^2 - \frac{1}{2} \right) = -\hat{\beta}^2 - \frac{1}{2}, \] (7)
where $\hat{\beta} = \frac{1}{M} \sum_{i=1}^{M} \hat{\beta}_i$ is the average of regression slopes in the least square linear regression on pairs $(j, \log_2 (\hat{\mu}_{j,i}))$ over the $M$ groups, $i = 1, 2, \ldots, M$.

The proof of Theorem III.1 is deferred to the appendix.

To find the optimal $\alpha$ and $p$ minimizing the asymptotic variance of $\hat{\mu}_{j,i}$, we take partial derivatives of $f(\alpha, p)$ with respect to $\alpha$ and $p$ and set them to 0. The optimal $\hat{\alpha}$ and $\hat{p}$ can be obtained by solving

$$\begin{align*}
\frac{\partial f(\alpha, p)}{\partial \alpha} &= -\frac{2p-1}{2p(1-p)} \alpha + \frac{1+p}{2(1-p)} - \frac{3}{2} = 0, \\
\frac{\partial f(\alpha, p)}{\partial p} &= \frac{\alpha (2-\alpha)}{2(1-p)^2} + \frac{\alpha^2 (2p-1)}{4p^2 (1-p)^2} = 0.
\end{align*}$$

Since $p \in (0, 1/2)$, and $\alpha \in [0, 1]$, we get the unique solution $p = 1 - \sqrt{2}/2 \approx 0.3$ and $\alpha = 2p \approx 0.6$. The Hessian matrix of $f(\alpha, p)$ is

$$\begin{bmatrix}
\frac{\partial^2 f(\alpha, p)}{\partial \alpha^2} & \frac{\partial^2 f(\alpha, p)}{\partial \alpha \partial p} \\
\frac{\partial^2 f(\alpha, p)}{\partial \alpha \partial p} & \frac{\partial^2 f(\alpha, p)}{\partial p^2}
\end{bmatrix} = \begin{bmatrix}
-\frac{2p-1}{2p(1-p)} & 2p^2 - 2p^2 + \alpha (2p-1) \\
2p^2 - 2p^2 + \alpha (2p-1) & 2p^2 (1-p) + \alpha^2 (2p-1)^2
\end{bmatrix}.
$$

Since $-\frac{2p-1}{2p(1-p)} > 0$ and the determinant is $5.66 > 0$ when $p = 1 - \sqrt{2}/2 \approx 0.3$ and $\alpha = 2p \approx 0.6$, the above Hessian matrix is positive definite. Therefore, $\hat{p} = 1 - \sqrt{2}/2$ and $\hat{\alpha} = 2 - \sqrt{2}$ provide the global minima of $f(\alpha, p)$, minimizing also the asymptotic variance of $\hat{\mu}_{j,i}$. In comparing these optimal $\hat{\alpha} \approx 0.6$ and $\hat{p} \approx 0.3$ with $\alpha = 0.6$ and $p = 1/3$ from the Gastwirth estimator, curiously, we find that the optimal general trimean estimator is very close to the Gastwirth estimator.

B. General Trimean of the Logarithm of Mid-energy (GTLME) Method

Previously discussed the GTME method calculates the general trimean estimator of the mid-energy first and then takes the logarithm. In this section, we will calculate the general trimean estimator of the logged mid-energies at each subgroup from level $j$. The following describes the general trimean of the logarithm of mid-energy, the GTLME method.

Theorem III.2. Let $\hat{\mu}_{j,i}$ be the general trimean estimator based on

$$\log (D_{j,i}), \log (D_{j,i+M}), \ldots, \log (D_{j,(N/2-M+i)}) := I(i,j)$$

where $I(i,j)$ is the $i$th group of logged mid-energies” at level $j$ in a $J$-level NDWT of a fBm of size $N$ with Hurst exponent $H$, $1 \leq i \leq M$ and $1 \leq j \leq J$. Then, the asymptotic distribution of $\hat{\mu}_{j,i}$ is normal,

$$\hat{\mu}_{j,i} \approx \mathcal{N} \left( c(\alpha, p) + \log (\lambda_j), \frac{2M}{N} f(\alpha, p) \right),$$

where

$$c(\alpha, p) = \frac{\alpha}{2} \log \left( \frac{1}{1-p} \cdot \log \frac{1}{p} \right) + (1 - \alpha) \log (2),$$

and

$$f(\alpha, p) = \frac{\alpha^2}{4g_1(p)} + \frac{\alpha (1 - \alpha)}{2g_2(p)} + \frac{(1 - \alpha)^2}{(\log 2)^2},$$

$g_1(p)$ and $g_2(p)$ are two functions of $p$.

$$\lambda_j = \sigma^2 \cdot 2^{-(2H+1)},$$

and $\sigma^2$ is the variance of wavelet coefficients from level 0. The Hurst exponent can be estimated as

$$\hat{H} = \frac{1}{M} \sum_{i=1}^{M} \hat{H}_i = \frac{1}{M} \sum_{i=1}^{M} \left( \frac{\hat{\beta}_i}{2 \log 2} - \frac{1}{2} \right) = -\frac{1}{2 \log 2} \frac{\hat{\beta} - 1}{2},$$

where $\hat{\beta} = \frac{1}{M} \sum_{i=1}^{M} \hat{\beta}_i$ is the average of regression slopes in the least square linear regressions on pairs $(j, \hat{\mu}_{j,i})$ over the $M$ groups, $i = 1, 2, \ldots, M$.

The proof of Theorem III.2 is provided in the appendix. Similarly, as for the GTME, the optimal $\alpha$ and $p$ which minimize the asymptotic variance of $\hat{\mu}_{j,i}$ can be obtained by solving

$$\frac{\partial f(\alpha, p)}{\partial \alpha} = 0, \quad \text{and} \quad \frac{\partial f(\alpha, p)}{\partial p} = 0.$$  

From the first equation in (11) it can be derived that

$$\alpha = \frac{2}{\log (2)^2} \frac{1}{g_1(p) - g_2(p) + \frac{2}{(\log 2)^2}}.$$  

The second equation in (11) cannot be simplified to a finite form. As an illustration, we plot the $f(\alpha, p)$ with $p$ ranging from 0 to 0.5 and $\alpha$ being a function of $p$. The plot of $\alpha$ against $p$ is also shown in Figure 1. Numerical calculation gives $\hat{p} = 0.24$ and $\hat{\alpha} = 0.5965$. These optimal parameters are close to $\alpha = 0.5$ and $p = 0.25$ in the Tukey’s trimean estimator, but put some more weight on the median.

C. Special Cases: Tukey’s Trimean and Gastwirth Estimators

The Tukey’s trimean of the mid-energy (TTME) method and Gastwirth of the mid-energy (GME) method are described in the following Lemma.

Lemma 2. Let $\hat{\mu}_{j,i}^T$ and $\hat{\mu}_{j,i}^G$ be the Tukey’s trimean and Gastwirth estimators based on $D(i,j)$ defined in the Theorem.
Then the asymptotic distributions of $\hat{\mu}_{j,i}^T$ and $\hat{\mu}_{j,i}^G$ are normal:

$$\hat{\mu}_{j,i}^T \sim N\left(c_1\lambda_j, \frac{5M}{3N} \lambda_j^2\right),$$  

(12)

$$\hat{\mu}_{j,i}^G \sim N\left(c_2\lambda_j, \frac{1.67M}{N} \lambda_j^2\right),$$  

(13)

where $c_1$ and $c_2$ are constants independent of $j$, $\lambda_j = \sigma^2 \cdot 2^{-(2H+1)j}$, and $\sigma^2$ is the variance of wavelet coefficients from level 0. The Hurst exponent can be estimated as

$$\hat{H}^T = -\frac{3}{2} \log 2 - \frac{\hat{\beta}^T}{2} - \frac{1}{2}, \quad \text{and} \quad \hat{H}^G = -\frac{3}{2} \log 2 - \frac{\hat{\beta}^G}{2} - \frac{1}{2},$$  

(14)

where $\hat{\beta}^T = \frac{1}{M} \sum_{i=1}^{M} \hat{\beta}_{i}^T$ and $\hat{\beta}^G = \frac{1}{M} \sum_{i=1}^{M} \hat{\beta}_{i}^G$ are the averages of regression slopes in the least square linear regression on pairs $(j, \log_2 (\hat{\mu}_{j,i}^T))$ and pairs $(j, \log_2 (\hat{\mu}_{j,i}^G))$ over the $M$ groups, respectively ($i = 1, 2, ..., M$).

The following Lemma describes the Tukey’s trimean (TTLME) and Gastwirth (GLME) of the logarithm of mid-energy method.

**Lemma 3.** Let $\hat{\mu}_{j,i}^T$ and $\hat{\mu}_{j,i}^G$ be the Tukey’s trimean estimator and Gastwirth estimator based on $L(i, j)$ defined in the Theorem III.2. The asymptotic distributions of $\hat{\mu}_{j,i}^T$ and $\hat{\mu}_{j,i}^G$ are normal,

$$\hat{\mu}_{j,i}^T \sim N\left(- (2H+1) \log 2j + c_3, V_1\right),$$  

(15)

$$\hat{\mu}_{j,i}^G \sim N\left(- (2H+1) \log 2j + c_4, V_2\right),$$  

(16)

where $c_3, V_1$, $c_4$ and $V_2$ are constants independent of level $j$. The Hurst exponent can be estimated as

$$\hat{H}^T = -\frac{1}{2 \log 2} \hat{\beta}^T - \frac{1}{2}, \quad \text{and} \quad \hat{H}^G = -\frac{1}{2 \log 2} \hat{\beta}^G - \frac{1}{2},$$  

(17)

where $\hat{\beta}^T = \frac{1}{M} \sum_{i=1}^{M} \hat{\beta}_{i}^T$ and $\hat{\beta}^G = \frac{1}{M} \sum_{i=1}^{M} \hat{\beta}_{i}^G$ are the averages of regression slopes in the least square linear regression on pairs $(j, \hat{\mu}_{j,i}^T)$ and pairs $(j, \hat{\mu}_{j,i}^G), i = 1, 2, ..., M$, over the $M$ groups.

The proofs of Lemma 2 and Lemma 3 are provided in the appendix.

**IV. Simulation**

We simulate fractional Brownian motion (fBm) signals of sizes $N = 2^{10}, N = 2^{11}$, and $N = 2^{12}$ with Hurst exponent $H = 0.3, 0.5, 0.7, 0.8, 0.9$, respectively. NDWT of depth $J = 10$ using Pollen wavelets with angles $\pi/6$ (Daubechies 2), $\pi/4$, $\pi/3$, and $\pi/2$ (Haar) are performed on each simulated signal to obtain wavelet coefficients. Pollen generates a family possessing continuum many wavelet bases of various degrees of regularity [19]. Special cases of Pollen’s representation for $\pi/6$ and $\pi/2$ give Daubechies 2 filter and Haar filter, respectively. Figure 2 depicts scaling and wavelet functions for $\pi/4$.

The proposed methods (with 6 variations) are applied on the NDWT coefficients to estimate Hurst exponent $H$. Coefficients on each level are divided into eight groups ($M = 8$) for all proposed methods, and we use wavelet coefficients from levels 4 to 10 for the least square linear regression. The estimation performance of the proposed methods are compared to four other existing methods: Veitch & Abry (VA) method, Soltani, Simard, & Boichu (SSB) method, MEDL method, and MEDLA method. The GTME and GTLME methods are based on the optimal parameters to minimize the variances. Estimation performances are compared in terms of mean, variance, and mean square error (MSE) based on 300 repetitions for each case.

The proposed methods perform the best using Haar wavelet (Pollen wavelets with angle $\pi/2$), and the simulation results are shown in Table I to Table III for fBm of sizes $N = 2^{10}, N = 2^{11}$, and $N = 2^{12}$, respectively. Similar results are obtained for other wavelets. For each $H$ (corresponding to each row in the table), the smallest variances and MSES are highlighted in bold. From simulations results, at least one of our 6 variations outperforms MEDL and MEDLA for all $H$ and fBm of all three sizes. Compared with VA and SSB methods, our methods yield significantly smaller variances and MSES when $H > 0.5$ for fBm of all three sizes. When $H = 0.3$ and 0.5, our methods are still comparable to VA and SSB. Although the performances of our 6 variations are very similar regarding to variances and MSES, the TTMME method based on Tukey’s trimean estimator of the mid-energy has the best performance among all of them. As expected, the variances of GTME based on the optimal parameters are smaller than or equal to those of GME and TTMME methods in most cases, and similarly, in most cases the optimized GTLME method is superior to other logged mid-energy methods TTLME and GLME with respect to variances. However, this superiority is not significant, since the variances of all six proposed methods are close to each other.

![Fig. 2: Pollen scaling and wavelet functions for π/4](image)

**V. Application**

In this section, we apply the proposed methodology to high frequency measurements in pupillary response behavior (PRB). PRB refers to changes in pupil diameter in response to simple or complex stimuli. Participants in this study included 24 older adults, solicited from the patient pool of the Bascom Palmer Eye Institute of the University of Miami School of Medicine. Participants were selected on the basis of having either no ocular disease or only Age-related Macular Degeneration (AMD), as assessed by patient history and clinical testing. Participants were selected based on the diagnosis of AMD and their best eye near distance (40 cm) visual acuity.
Participants were assigned to 3 groups (one control group and 2 experimental groups). The control group is a set of individuals with healthy, unaffected vision and no evidence of any ocular disease or trauma. Individuals in 2 experimental groups had varying visual acuity and were diagnosed with AMD. The number of participants is 6 in control group, 12 in group I, and 6 in group II. The data analysis process consists of first cleaning the data by removing the blink and equipment artifacts and then segmenting the data stream for each participant into equal length pieces of 2048 observations. The number of 2048 length data sets that were extracted from the collective data sets of the individuals within each group are shown in Table IV.

Traditional statistical methods have not been successfully used for examining the PRB of older adults or individuals with visual impairments. Researchers have utilized simple statistical methods for analyzing PRB, for example, comparing the relative mean or variance of pupil size deviation in response to stimuli; some sophisticated techniques have also been utilized, like power, frequency and spectral analysis using mathematical tools. But most of them failed to characterize the underlying patterns within time series PRB data. Wavelet analysis to estimate the Hurst exponent of the high-frequency, time series physiological data is a useful tool for detecting these hidden patterns and differentiating individuals based on these unique patterns in their physiological behavior.

Table V provides descriptive statistics of the estimated Hurst exponent $H$ in each group using our proposed methods and four standard methods to compare with. As can be seen, there are clear monotonic trends for $H$ across the participant groups. The control group exhibited the smallest value for $H$, followed by group I and group II, and those monotonic trends are

| TABLE I: Simulation Results for $N = 2^{10}$ fBm using Haar wavelet (300 Replications) |
|----------------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $H$ | VA    | SSB   | MEDL  | MEDLA  | TTME  | GME   | TTLME | GLME  | GTME  | GTLME |
|------|-------|-------|-------|--------|-------|-------|-------|-------|-------|-------|
| 0.3  | 0.2429| 0.2429| 0.2368| 0.2365 | 0.2397| 0.2393| 0.2393| 0.2390| 0.2395| 0.2396|
| 0.5  | 0.4482| 0.4944| 0.4800| 0.4821 | 0.4893| 0.4870| 0.4872| 0.4861| 0.4882| 0.4876|
| 0.7  | 0.5209| 0.7232| 0.6996| 0.7002 | 0.7094| 0.7072| 0.7067| 0.7058| 0.7085| 0.7071|
| 0.8  | 0.5171| 0.8304| 0.7999| 0.7991 | 0.8050| 0.8020| 0.8031| 0.8008| 0.8037| 0.8044|
| 0.9  | 0.4976| 0.9410| 0.9020| 0.8951 | 0.9016| 0.9004| 0.9024| 0.9009| 0.9012| 0.9032|

| TABLE II: Simulation Results for $N = 2^{11}$ fBm using Haar wavelet (300 Replications) |
|----------------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $H$ | VA    | SSB   | MEDL  | MEDLA  | TTME  | GME   | TTLME | GLME  | GTME  | GTLME |
|------|-------|-------|-------|--------|-------|-------|-------|-------|-------|-------|
| 0.3  | 0.2448| 0.2420| 0.2385| 0.2382 | 0.2393| 0.2389| 0.2392| 0.2390| 0.2392| 0.2394|
| 0.5  | 0.4548| 0.4822| 0.4721| 0.4725 | 0.4755| 0.4745| 0.4743| 0.4740| 0.4749| 0.4744|
| 0.7  | 0.5216| 0.7125| 0.6917| 0.6931 | 0.6977| 0.6963| 0.6959| 0.6954| 0.6969| 0.6963|
| 0.8  | 0.5224| 0.8201| 0.7937| 0.7933 | 0.7986| 0.7972| 0.7979| 0.7970| 0.7979| 0.7985|
| 0.9  | 0.5028| 0.9301| 0.8988| 0.8922 | 0.8958| 0.8957| 0.8976| 0.8965| 0.8958| 0.8982|

### Variances

| $H$ | Variances |
|------|------------|
| 0.3  | 0.0024     | 0.0024 | 0.0025 | 0.0022 | 0.0020 | 0.0021 | 0.0022 | 0.0022 | 0.0022 |
| 0.5  | 0.0018     | 0.0034 | 0.0035 | 0.0030 | 0.0030 | 0.0031 | 0.0031 | 0.0031 | 0.0030 |
| 0.7  | 0.0048     | 0.0055 | 0.0049 | 0.0038 | 0.0036 | 0.0036 | 0.0038 | 0.0037 | 0.0036 |
| 0.8  | 0.0067     | 0.0071 | 0.0061 | 0.0044 | 0.0041 | 0.0044 | 0.0047 | 0.0047 | 0.0047 |
| 0.9  | 0.0064     | 0.0085 | 0.0066 | 0.0047 | 0.0043 | 0.0043 | 0.0047 | 0.0046 | 0.0042 |

### MSEs

| $H$ | MSEs |
|------|------|
| 0.3  | 0.0057 | 0.0056 | 0.0065 | 0.0062 | 0.0056 | 0.0058 | 0.0059 | 0.0059 | 0.0057 | 0.0058 |
| 0.5  | 0.0044 | 0.0034 | 0.0038 | 0.0034 | 0.0031 | 0.0032 | 0.0032 | 0.0033 | 0.0031 | 0.0032 |
| 0.7  | 0.0368 | 0.0060 | 0.0061 | 0.0044 | 0.0042 | 0.0044 | 0.0047 | 0.0047 | 0.0047 | 0.0047 |
| 0.8  | 0.0867 | 0.0080 | 0.0061 | 0.0044 | 0.0041 | 0.0043 | 0.0047 | 0.0046 | 0.0042 | 0.0047 |
| 0.9  | 0.1683 | 0.0101 | 0.0066 | 0.0047 | 0.0043 | 0.0043 | 0.0047 | 0.0046 | 0.0042 | 0.0047 |
presented in both the mean and median values. In fact, signals with smaller Hurst exponent $H$ tend to be more disordered and unsystematic, therefore the monotonic trends indicating that individuals with higher severity of visual impairment have less disordered pupil diameter signals. Similar results for original data without being pre-processed are shown in Table VI. No monotonic trends of $\hat{H}$ across groups have been observed in original data, indicating the blinks indeed add noise to the data.

Like in many human-subject studies, the limited number of participants is a major disadvantage, but in PRB data set, each subject has enough measurements to segment into multiple pieces with a length of 2048 observations. This induces dependence between the runs, and hierarchical models accommodating for the subject induced dependence are needed. If we denote $i$ to be the group index where the piece of observations is from, with $i = 0$ for control group, $i = 1$ for group I, $i = 2$ for group II, and $n_i$ as the number of pieces generated from subject $j \ (j=1,2,..,24)$, the estimated Hurst exponent $\hat{H}_{ijk}$ for the $k$th piece of subject $j$ nested in group $i$ can be expressed in the following model:

$$\hat{H}_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}, \quad (18)$$

where $\mu$ is the overall mean, $\alpha_i$ is the effect for $i$th group, $\beta_j$ is the effect for $j$th participant within $i$th group, and $\epsilon_{ijk}$ is the random error. The objective is to classify the groups based on the estimated Hurst exponent for a given pupil diameter data. In avoid of dependency between data sets due to subject effects, the estimated $\hat{\beta}_j$ is first subtracted from $\hat{H}_{ijk}$, and multinominal logistic regression model is fitted on the data.

$$\left\{ \hat{H}_{ijk} - \hat{\beta}_j, i = 0, 1, 2, j = 1, 2, ..., n_j \right\}$$

To test the model performance, we randomly choose 80% of the data points to form a training set, and the remaining forms the testing set. Model is developed on the training set and applied on the testing set; misclassification rate is reported in Table VII.
### TABLE IV: Group characterization summary

| Group | N | Visual Acuity | AMD | Number of preprocessed data sets |
|-------|---|---------------|-----|----------------------------------|
| Control | 6 | 20/20-20/40 | No  | 49                              |
| I      | 12 | 20/20-20/100 | Yes | 184                             |
| II     | 6  | 20/100       | Yes | 170                             |

Note. N represents the number of individuals in the group; Visual Acuity represents the range of Snellen acuity scores for the individuals in the given group, AMD represents whether the individuals were diagnosed with age-related macular degeneration or not, and Number of cleaned data sets shows the number of 2048 length data sets extracted from individuals within each group.

### TABLE V: Descriptive Statistics Group Summary (Blinks removed)

| Methods | Group | Veitch and Abry | Soltani | MEDL | MEDLA | TTME | GME | TTLME | GLME | GTME | GTLME |
|---------|-------|-----------------|---------|------|------|------|-----|-------|------|------|-------|
| Mean of $\hat{H}$ | Control | 0.1495 | 0.4103 | 0.2741 | 0.2501 | 0.2416 | 0.2505 | 0.2537 | 0.2559 | 0.2467 | 0.2530 |
|          | I    | 0.1855 | 0.4210 | 0.2733 | 0.2277 | 0.2638 | 0.2686 | 0.2716 | 0.2738 | 0.2666 | 0.2705 |
|          | II   | 0.1682 | 0.5464 | 0.3597 | 0.3202 | 0.2984 | 0.3116 | 0.3257 | 0.3231 | 0.3057 | 0.3286 |
| Median of $\hat{H}$ | Control | 0.1474 | 0.4088 | 0.2869 | 0.2596 | 0.2535 | 0.2609 | 0.2625 | 0.2683 | 0.2577 | 0.2615 |
|          | I    | 0.2073 | 0.4331 | 0.3033 | 0.2951 | 0.2938 | 0.2990 | 0.3054 | 0.3038 | 0.2976 | 0.3042 |
|          | II   | 0.1814 | 0.5436 | 0.3514 | 0.3240 | 0.2973 | 0.3127 | 0.3261 | 0.3313 | 0.3058 | 0.3254 |
| Variance of $\hat{H}$ | Control | 0.0084 | 0.0195 | 0.0086 | 0.0091 | 0.0094 | 0.0092 | 0.0094 | 0.0095 | 0.0092 | 0.0092 |
|          | I    | 0.0120 | 0.0421 | 0.0213 | 0.0206 | 0.0191 | 0.0196 | 0.0209 | 0.0204 | 0.0194 | 0.0198 |
|          | II   | 0.0114 | 0.0373 | 0.0428 | 0.0066 | 0.0061 | 0.0060 | 0.0087 | 0.0065 | 0.0060 | 0.0102 |

### TABLE VI: Descriptive Statistics Group Summary (Blinks not removed)

| Methods | Group | Veitch and Abry | Soltani | MEDL | MEDLA | TTME | GME | TTLME | GLME | GTME | GTLME |
|---------|-------|-----------------|---------|------|------|------|-----|-------|------|------|-------|
| Mean of $\hat{H}$ | Control | 0.1892 | 0.4627 | 0.3582 | 0.3497 | 0.3394 | 0.3498 | 0.3473 | 0.3498 | 0.3459 | 0.3446 |
|          | I    | 0.1812 | 0.4272 | 0.2860 | 0.2918 | 0.2858 | 0.2900 | 0.2900 | 0.2923 | 0.2891 | 0.2888 |
|          | II   | 0.1565 | 0.5510 | 0.3672 | 0.3295 | 0.3113 | 0.3262 | 0.3373 | 0.3360 | 0.3201 | 0.3416 |
| Median of $\hat{H}$ | Control | 0.2144 | 0.5194 | 0.3502 | 0.3574 | 0.3464 | 0.3559 | 0.3513 | 0.3542 | 0.3510 | 0.3508 |
|          | I    | 0.1867 | 0.4467 | 0.3072 | 0.3091 | 0.2943 | 0.2981 | 0.3100 | 0.3055 | 0.3014 | 0.3080 |
|          | II   | 0.1758 | 0.5528 | 0.3568 | 0.3279 | 0.2999 | 0.3219 | 0.3330 | 0.3345 | 0.3104 | 0.3320 |
| Variance of $\hat{H}$ | Control | 0.0183 | 0.0311 | 0.0294 | 0.0288 | 0.0258 | 0.0270 | 0.0253 | 0.0261 | 0.0266 | 0.0246 |
|          | I    | 0.0184 | 0.0449 | 0.0245 | 0.0245 | 0.0250 | 0.0251 | 0.0244 | 0.0248 | 0.0252 | 0.0242 |
|          | III  | 0.0212 | 0.0363 | 0.0420 | 0.0082 | 0.0098 | 0.0089 | 0.0102 | 0.0083 | 0.0093 | 0.0132 |

### TABLE VII: Classification error

| Methods | Group | Veitch and Abry | Soltani | MEDL | MEDLA | TTME | GME | TTLME | GLME | GTME | GTLME |
|---------|-------|-----------------|---------|------|------|------|-----|-------|------|------|-------|
| Original data | 0.3837 | 0.3721 | 0.3605 | 0.3372 | 0.3721 | 0.3721 | 0.3721 | 0.3721 | 0.3721 | 0.3837 |
| Blinks removed | 0.2840 | 0.3210 | 0.2469 | 0.2963 | 0.2469 | 0.2469 | 0.2469 | 0.2469 | 0.2469 | 0.2716 |

The proposed methods have been applied to a real pupillary response behavior (PRB) data set to extract meaningful characteristics from the PRB data of older adults with and without visual impairment. Estimated Hurst exponents base on wavelet analysis capture the unique pattern of a data signal that cannot be represented by the the trends and traditional statistical summaries of the signal in that the magnitude of the pupil diameter depends on ambient light, not on the inherent eye function. Our proposed methods helped to detect those unique patterns and differentiate individuals based on the estimated Hurst parameters $\hat{H}$. Monotonic trends have been found for $\hat{H}$ across the participant groups, and individuals with higher severity of visual impairment have more regular pupil diameter signals. This increase of regularity with increase of the degree of pathology is common for many other biometric signals: EEG, EKG, high frequency protein mass-spectra, high resolution medical images of tissue, to list a few.

### APPENDIX A: TECHNICAL PROOFS.

**Proof of Theorem III.1.**

**Proof.** A single wavelet coefficient in a non-decimated wavelet transform of a fBm of size $N$ with Hurst exponent $H$ is to variances and MSEs, the TTME method based on Tukey's trimean estimator of the mid-energy has the best performance among all of them.

The proposed methods have been applied to a real pupillary response behavior (PRB) data set to extract meaningful characteristics from the PRB data of older adults with and without visual impairment. Estimated Hurst exponents base on wavelet analysis capture the unique pattern of a data signal that cannot be represented by the the trends and traditional statistical summaries of the signal in that the magnitude of the pupil diameter depends on ambient light, not on the inherent eye function. Our proposed methods helped to detect those unique patterns and differentiate individuals based on the estimated Hurst parameters $\hat{H}$. Monotonic trends have been found for $\hat{H}$ across the participant groups, and individuals with higher severity of visual impairment have more regular pupil diameter signals. This increase of regularity with increase of the degree of pathology is common for many other biometric signals: EEG, EKG, high frequency protein mass-spectra, high resolution medical images of tissue, to list a few.
The pdf and cdf of $y_j$ can be readily shown that $D_j,k$ has exponential distribution with scale parameter $\lambda_j = \sigma^2 \cdot 2^{-(2H+1)j}$, i.e.,
\[
 f(D_j,k) = \lambda_j^{-1}e^{-\lambda_j^{-1}D_j,k}, \text{ for any } k = 1, \ldots, N/2.
\]

Therefore the asymptotic distribution of $\hat{\mu}_{j,i}$ is normal with mean
\[
 \mathbb{E}(\hat{\mu}_{j,i}) = A \cdot \xi = \left(\frac{\alpha}{2} \log \left(\frac{1}{p(1-p)}\right) + (1 - \alpha) \log 2\right) \lambda_j \triangleq c(\alpha, \lambda_j),
\]
and variance
\[
 \text{Var}(\hat{\mu}_{j,i}) = 2M \frac{\alpha^2}{N} \text{Var}(\hat{D}_{j,i}) = 2M \frac{\alpha(1 - 2p)(\alpha - 4p)}{4p(1-p)} + 1 \lambda_j^2 \triangleq 2M \frac{\alpha}{N} f(\alpha, \lambda_j)^2.
\]

**Proof of Theorem III.2.**

*Proof.* We have stated that $D_{j,k} \sim \text{Exp}\left(\lambda_j^{-1}\right)$ with scale parameter $\lambda_j = \sigma^2 \cdot 2^{-(2H+1)j}$, so that
\[
 f(D_{j,k}) = \lambda_j^{-1}e^{-\lambda_j^{-1}D_{j,k}}, \text{ for any } k = 1, \ldots, N/2.
\]

Let $y_{j,k} = \log(D_{j,k})$ for any $j = 1, \ldots, J$ and $k = 1, \ldots, N/2$. The pdf and cdf of $y_{j,k}$ are
\[
 f(y_{j,k}) = \lambda_j^{-1}e^{-\lambda_j^{-1}y_{j,k}},
\]
and
\[
 F(y_{j,k}) = 1 - e^{-\lambda_j^{-1}e^{y_{j,k}}}. \quad \text{(The p-quantile can be obtained by solving } F(y_p) = 1 - e^{-\lambda_j^{-1}e^{y_{j,k}}} = p, \text{ and } y_p = \log(-\lambda_j \log(1-p))).
\]
and variance
\[
\operatorname{Var} (\hat{\mu}^T_{j,i} ) = \frac{2M}{N} A_T \Sigma_T A^T_T = \frac{5M}{3N} \lambda_j^2.
\]

When applying Gastwirth estimator \( \hat{\mu}^G_{j,i} \) on \( \{D_{j,i}, D_{j,i+M}, \ldots, D_{j,(N/2−M+i)}\} \), following the derivation in Section II-B, we have
\[
\xi_G = \begin{bmatrix}
\log (\log (\frac{3}{4})) + \log (2\lambda_j) \\
\log (\lambda_j) + \log (2\lambda_j) \\
\log (2\lambda_j) + \log (2\lambda_j)
\end{bmatrix},
\]
and
\[
\Sigma_G = \begin{bmatrix}
\frac{1}{2\log (\frac{3}{4})} & \frac{1}{2\log (\frac{3}{4})} & \frac{1}{2\log (\frac{3}{4})} \\
\frac{1}{2\log (\frac{3}{4})} & \frac{1}{2\log (\frac{3}{4})} & \frac{1}{2\log (\frac{3}{4})} \\
\frac{1}{2\log (\frac{3}{4})} & \frac{1}{2\log (\frac{3}{4})} & \frac{1}{2\log (\frac{3}{4})}
\end{bmatrix},
\]
therefore, the asymptotic distribution of \( \hat{\mu}^G_{j,i} \) is normal with mean
\[
\mathbb{E} (\hat{\mu}^G_{j,i} ) = A_G \cdot \xi_G = \begin{bmatrix}
0.3 \times \log (\log (\frac{9}{2})) + 0.4 \times \log (2) \lambda_j \\
0.3 \times \log (\log (\log (4))) + 0.4 \times \log (\log (4)) + 0.3 \times \log (\log (9)) \\
\end{bmatrix} \triangleq c_2 \lambda_j,
\]
and variance
\[
\operatorname{Var} (\hat{\mu}^G_{j,i} ) = \frac{2M}{N} A_G \Sigma_G A^T_G = \frac{1.67M}{N} \lambda_j^2.
\]

Proof of Lemma 3.

Proof. When applying Tukey’s trimean estimator \( \hat{\mu}^T_{j,i} \) on \( \{D_{j,i}, D_{j,i+M}, \ldots, D_{j,(N/2−M+i)}\} \), following the derivation in Section II-A, we have
\[
\xi_T = \begin{bmatrix}
\log (\log (\frac{16}{9})) + \log (2\lambda_j) \\
\log (\log (4)) + \log (2\lambda_j) \\
\log (16) + \log (2\lambda_j)
\end{bmatrix},
\]
and
\[
\Sigma_T = \begin{bmatrix}
\frac{1}{3\log (\frac{3}{4})} & \frac{1}{3\log (\frac{3}{4})} & \frac{1}{3\log (\frac{3}{4})} \\
\frac{1}{3\log (\frac{3}{4})} & \frac{1}{3\log (\frac{3}{4})} & \frac{1}{3\log (\frac{3}{4})} \\
\frac{1}{3\log (\frac{3}{4})} & \frac{1}{3\log (\frac{3}{4})} & \frac{1}{3\log (\frac{3}{4})}
\end{bmatrix},
\]
therefore, the asymptotic distribution of \( \hat{\mu}^T_{j,i} \) is normal with mean
\[
\mathbb{E} (\hat{\mu}^T_{j,i} ) = A_T \cdot \xi_T = -(2H+1) \log 2j - \log 2 + \log \sigma^2 + \frac{1}{4} \log \left( \log \left( \frac{16}{9} \right) \cdot \log (16) \right) + \frac{1}{2} \log (\log (4)) \triangleq -(2H+1) \log 2j + c_3
\]
and variance
\[
\operatorname{Var} (\hat{\mu}^T_{j,i} ) = \frac{2M}{N} A_T \Sigma_T A^T_T = \frac{M}{2N} \left( \frac{1}{12\log (\frac{3}{4})} + \frac{1}{3\log (\frac{3}{4})} + \frac{1}{6\log (\frac{3}{4})} \right) + \frac{1}{\log (\frac{3}{4})} \cdot \frac{1}{\log (\frac{3}{4})} \cdot \frac{3}{4\log (\frac{3}{4})} \triangleq V_1.
\]
Chen Feng is a Ph.D. student in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology. She received her B.S. degree in Applied Mathematics from Xi’an Jiaotong University (XJTU), Xi’an, China, in 2015, and Her research interests include wavelets, big data methodology, biostatistics, clinical trials, and statistical modeling in healthcare.

Brani Vidakovic is a Professor of Statistics at the School of Industrial and Systems Engineering and Department Biomedical Engineering at Georgia Institute of Technology and Emory University. His research focuses on biostatistics, Bayesian methodology, and statistical modeling in multiscale domains with applications in bioengineering, medicine and environmental sciences. Vidakovic received his bachelor and masters degrees in mathematics from University of Belgrade (Serbia) in 1978, and 1981, respectively, and Ph. D. in statistics at Purdue University (Indiana, USA) in 1992. He was an Assistant and Associate Professor at Institute of Statistics and Decision Sciences at Duke University form 1992 to 2000 prior to joining Georgia Institute of Technology. Vidakovic is a Fellow of American Statistical Association and elected member of International Statistical Institute.

REFERENCES

[1] K. P. Moloney, J. A. Jacko, B. Vidakovic, F. Sainfort, V. K. Leonard, and B. Shi, “Leveraging data complexity: Pupillary behavior of older adults with visual impairment during hci,” ACM Trans. Comput.-Hum. Interact., vol. 13, no. 3, pp. 376–402, Sep. 2006. [Online]. Available: http://doi.acm.org/10.1145/1183456.1183460

[2] M. Kang and B. Vidakovic, “MEDL and MEDLA: Methods for Assessment of Scaling by Medians of Log-Squared Nondecimated Wavelet Coefficients,” ArXiv e-prints, Mar. 2017.

[3] K. AN, “Wienersche spiralen und einige andere interessante kurven im hilbertschen raum,” Acad Sci USSR (NS), vol. 26, pp. 115–118, 1940.

[4] B. B. Mandelbrot and J. W. Van Ness, “Fractional brownian motions, fractional noises and applications,” SIAM review, vol. 10, no. 4, pp. 422–437, 1968.

[5] P. Abry, P. Flandrin, M. S. Taqqu, D. Veitch et al., “Wavelets for the analysis, estimation and synthesis of scaling data,” Self-similar network traffic and performance evaluation, pp. 39–88, 2000.

[6] P. Abry, P. Gonçalves, and P. Flandrin, “Wavelets, spectrum analysis and 1/f processes,” LECTURE NOTES IN STATISTICS-NEW YORK-SPRINGER VERLAG-, pp. 15–15, 1995.

[7] P. Abry, P. Gonçalves, and J. L. Véhel, Scaling, fractals and wavelets, John Wiley & Sons, 2013.

[8] S. Soltani, P. Simard, and D. Boichu, “Estimation of the self-similarity parameter using the wavelet transform,” Signal Processing, vol. 84, no. 1, pp. 117–123, 2004.

[9] H. Shen, Z. Zhu, and T. C. Lee, “Robust estimation of the self-similarity parameter in network traffic using wavelet transform,” Signal Processing, vol. 87, no. 9, pp. 2111–2124, 2007.

[10] J. W. Tukey, “Exploratory data analysis,” 1977.

[11] D. F. Andrews and F. R. Hampel, Robust estimates of location: Survey and advances. Princeton University Press, 2015.

[12] G. J. L., “on robust procedures,” Journal of the American Statistical Association, vol. 61, p. 929, 1966.

[13] G. J. L. and C. M., “small sample behavior of some robust linear estimates of locations,” Journal of the American Statistical Association, vol. 65, p. 946, 1970.

[14] G. J. L. and R. H., “on robust linear estimators,” Annals of Mathematical Statistics, vol. 40, p. 24, 1969.

[15] A. DasGupta, Asymptotic theory of statistics and probability. Springer Science & Business Media, 2008.

[16] G. P. Nason and B. W. Silverman, “The stationary wavelet transform and some statistical applications,” LECTURE NOTES IN STATISTICS-NEW YORK-SPRINGER VERLAG-, pp. 281–281, 1995.

[17] B. Vidakovic, Statistical modeling by wavelets, John Wiley & Sons, 2009, vol. 503.

[18] D. B. Percival and A. T. Walden, Wavelet methods for time series analysis, Cambridge university press, 2006, vol. 4.

[19] D. Pollen, “Su i (2, f [z, 1/z]) for fa subfield of c,” Journal of the American Mathematical Society, pp. 611–624, 1990.