Name: ______________________________

This exam is closed book and consists of 7 questions. With the exception of question #1 (which is worth 24 points), each question is worth 12 points, for a total of 96 points (an additional 4 points will be awarded for writing your name in the space provided above!).

[1] Consider a time series $\mathbf{X} = [X_0, X_1]^T$ of length $N = 2$.

(a) Write down what the Haar DWT matrix $\mathcal{W}$ would be for $\mathbf{X}$.

(b) In terms of X_0 and X_1, what are the Haar DWT coefficients for $\mathbf{X}$?

(c) Using your answer to part (b), demonstrate explicitly that the energy in $\mathbf{X}$ is equal to the energy in its Haar DWT coefficients.

(d) Determine what the first level detail $\mathcal{D}_1$ is explicitly in terms of X_0 and X_1.

(e) Determine what the first level smooth $\mathcal{S}_1$ is explicitly in terms of X_0 and X_1.

(f) Using your answers to parts (d) and (e), demonstrate explicitly that $\mathcal{D}_1 + \mathcal{S}_1 = \mathbf{X}$

(g) What are the level $J_0 = 1$ Haar MODWT coefficients for $\mathbf{X}$?

(h) Using your answer to part (g), demonstrate explicitly that the energy in $\mathbf{X}$ is equal to the energy in its Haar MODWT coefficients (hint: reuse your answer to part (c)).

[2] Let L be the width of a Daubechies wavelet filter $\{h_l\}$, and assume that any L under consideration is always much smaller than the sample size N.

(a) As L increases, does the maximum level out to which we can form a DWT for a time series of sample size $N = 2^J$ increase, decrease or remain the same (circle one)?

(b) As L increases, does the number of wavelet coefficients in a level J_0 DWT increase, decrease or remain the same (circle one)?

(c) As L increases, does the number of boundary wavelet coefficients in a level J_0 DWT increase, decrease or remain the same (circle one)?

(d) As L increases, does an LA(L) DWT wavelet filter yield an approximation to an ideal high-pass filter that is better, worse or remains the same (circle one)?

(e) In comparison to an D(L) DWT wavelet filter, does an LA(L) DWT wavelet filter yield an approximation to an ideal high-pass filter that is better, worse or equivalent (circle one)?

(f) As L increases, does the nominal bandwidth (as measured by the '3 dB down' point) for the passband of a wavelet filter increase, decrease or remain the same (circle one)?

[3] Let $\mathbf{X}$ be an N dimensional vector containing a time series, and let $\widetilde{\mathbf{W}}_1$ and $\widetilde{\mathbf{V}}_1$ be its MODWT wavelet and scaling coefficients for unit scale. Recall that $\tilde{\mathcal{B}}_1$ and $\tilde{\mathcal{A}}_1$ are $N \times N$ matrices such that $\tilde{\mathcal{B}}_1 \mathbf{X} = \widetilde{\mathbf{W}}_1$ and $\tilde{\mathcal{A}}_1 \mathbf{X} = \widetilde{\mathbf{V}}_1$. Using the fact that $\tilde{\mathcal{B}}_1^T \tilde{\mathcal{B}}_1 + \tilde{\mathcal{A}}_1^T \tilde{\mathcal{A}}_1 = I_N$ (the $N \times N$ identity matrix), show that

$$\mathbf{X} = \widetilde{\mathcal{B}}_1^T \widetilde{\mathbf{W}}_1 + \widetilde{\mathcal{A}}_1^T \widetilde{\mathbf{V}}_1 \quad \text{and} \quad \|\mathbf{X}\|^2 = \|\widetilde{\mathbf{W}}_1\|^2 + \|\widetilde{\mathbf{V}}_1\|^2.$$

[4] Recall that the best basis algorithm consists of two basic steps. First, given a wavelet packet (WP) table to level J_0, we associate with each $\mathbf{W}_{j,n}$ a cost

$$M(\mathbf{W}_{j,n}) \equiv \sum_{t=0}^{N_j-1} m(|W_{j,n,t}|),$$

where $m(\cdot)$ is a real-valued cost function. Second, we define the best basis as the solution to

$$\min_{\mathcal{C}} \sum_{(j,n)\in\mathcal{C}} M(\mathbf{W}_{j,n});$$

i.e., we seek the orthonormal transform specified by $\mathcal{C}$ such that the cost summed over all doublets $(j,n) \in \mathcal{C}$ is minimized. Briefly discuss the advantages and disadvantages of using the cost functional $m(|W_{j,n,t}|) = |W_{j,n,t}|^2$ for selecting a best basis.

[5] In the first step of the matching pursuit algorithm, we project a time series $\mathbf{X}$ against a vector $\mathbf{d}$ that is an element of a large collection of normalized vectors called a dictionary. We then form an approximation to $\mathbf{X}$ defined by $\widehat{\mathbf{X}} = \langle \mathbf{X}, \mathbf{d} \rangle \mathbf{d}$ and a residual vector defined by $\mathbf{R} = \mathbf{X} - \widehat{\mathbf{X}}$. Show that $\langle \widehat{\mathbf{X}}, \mathbf{R} \rangle = 0$ and hence that

$$\|\mathbf{X}\|^2 = \|\widehat{\mathbf{X}}\|^2 + \|\mathbf{R}\|^2 = |\langle \mathbf{X}, \mathbf{d} \rangle|^2 + \|\mathbf{R}\|^2$$

(recall that $\|\mathbf{X}\|^2 = \langle \mathbf{X}, \mathbf{X} \rangle$).

[6] Let $\hat{\nu}_X^2(\tau_j)$ be the MODWT-based unbiased estimator of the wavelet variance $\nu_X^2(\tau_j)$. Large sample theory says that the variance of $\hat{\nu}_X^2(\tau_j)$ is approximately $2A_j/M_j$, where A_j is a constant, while M_j is the number of MODWT wavelet coefficients used to form $\hat{\nu}_X^2(\tau_j)$.

(a) Under the assumption that $\hat{\nu}_X^2(\tau_j)$ is distributed as $a\chi_\eta^2$, where χ_η^2 denotes a chi-square random variable with η degrees of freedom, determine expressions for a and η (the EDOF) in terms of $E\{\hat{\nu}_X^2(\tau_j)\}$ and $\operatorname{var}\{\hat{\nu}_X^2(\tau_j)\}$ by solving the two equations $E\{\hat{\nu}_X^2(\tau_j)\} = E\{a\chi_\eta^2\}$ and $\operatorname{var}\{\hat{\nu}_X^2(\tau_j)\} = \operatorname{var}\{a\chi_\eta^2\}$.

(b) Use the result you got in (a) along with the large sample approximations to the mean and variance of $\hat{\nu}_X^2(\tau_j)$ to develop an expression for η in terms of quantities defined above.

[7] The left-hand column of Figure 1 shows four plots of wavelet variance estimates $\hat{\nu}_X(\tau_j)$ versus scale τ_j (plots (a), (b), (c) and (d)), while the right-hand column shows four time series (plots (1), (2), (3) and (4)). The time series are not necessarily placed next to their corresponding $\hat{\nu}_X(\tau_j)$'s; i.e., the time series in plot (1) is not necessarily associated with the estimates in plot (a) and so forth. In the following questions you are asked to match up the time series with their $\hat{\nu}_X(\tau_j)$'s and to state the reasons for your choices.

(a) The $\hat{\nu}_X(\tau_j)$'s in plot (a) go with the time series in plot ____ (fill in) because:

(b) The $\hat{\nu}_X(\tau_j)$'s in plot (b) go with the time series in plot ____ (fill in) because:

(c) The $\hat{\nu}_X(\tau_j)$'s in plot (c) go with the time series in plot ____ (fill in) because:

(d) The $\hat{\nu}_X(\tau_j)$'s in plot (d) go with the time series in plot ____ (fill in) because:

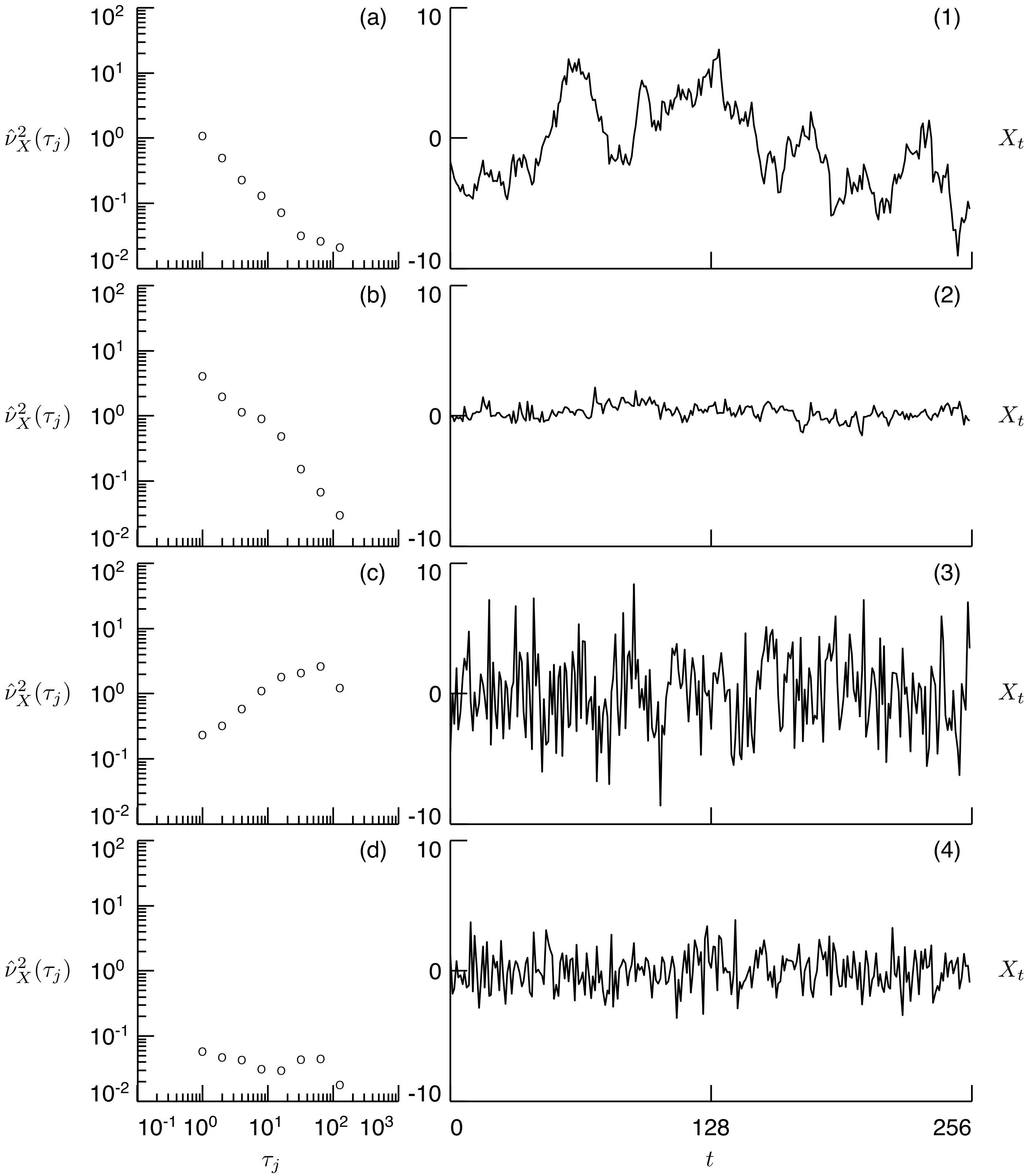

Figure 1.