François A. Marshall, *Member, IEEE*

I. Introduction

Starting and more detailed instructions for user execution of the code are provided in Section V. For example: the user can run the provided example code on `RStudio` simply by sourcing the `Run_Code.R` program under the subdirectory, `Setup/Run_Code`, of `Raw_Series_Plot_Epochs` or `Harmonic_Anharmonic`. Depending on the version of `R` being used, the code execution might result in the `RStudio` console asking the user questions about installing appropriate `R` libraries (something that could happen if the user is newly using `R`).

The data considered in the example code of the GitHub repository comprise a 30 kHz voltage series from a single microelectrode, and have been drawn from a complex clinical survey of neocortical seizure activity. Details of the survey can be found in [5], while the data are available at [4]. The electrode having unit identification index is considered from the Subject C5 study (refer to Table 1 in [5] for details).

To modify the code, the user refers to Section IV of this document, which itself presents: the code initialization; and specifically where in the code directories and files to find and modify the initialization from that which is provided in the starting example. Modifications can range from a change of plotting units and data files to changing the names of certain file directories - as specified in Section IV. The nature of these modifications is appreciably strict, but they do allow the code to be used for quite arbitrary univariate times-series datasets, and are especially useful for those wishing to perform basic harmonic and power spectral analyses on big time-series records (i.e., with record size exceeding 10^4). In addition, the record size could even be quite small (e.g., 100 data points), though beware that multitaper spectral estimators are liable to be less resolved in frequency and incur appreciable bias (refer to Chapters 4 and 5 in [1]).

Sections II and III display the starting setup and algorithms used to obtain the results of multitaper power-spectral and harmonic analyses. The section concludes with a list of abbreviations together with a list of references. All source code discussed in this document is included under the directories, `Raw_Series_Plot_Epochs` and `Harmonic_Anharmonic`, while all header files are included under `General_Headers`. All listed code functions are contained under the `Multitaper_Header.R` header file, itself found under the `General_Headers` directory. In the code files, remark that some of the variable names include "time" or "voltage" for the same reason that one uses the term, "spectral power" to describe spectral mass: historically, an important application of statistical signal processing has been in electrical engineering, where time series of voltage are predominantly the response outputs.

This work was supported by National Institute of Neurological Disorders and Stroke R01NS110669.

F. A. Marshall completed this work as Postdoctoral Associate (2020-2022) and Visiting Scholar (2022) with Mathematics and Statistics Boston University, 111 Cummington Mall #140C, Boston, MA 02215, United States (e-mail: francois.marshall@queensu.ca).

II. Starting Setup

This section details the probability model and quantities relevant to the algorithms in Section III.

Data

1) Sampling period:
   Δt.
2) Time series:
   $\mathbf{x} \in \mathbb{R}^N$.
3) Size of the global record from which $\mathbf{x}$ has been drawn:
   $J_{PRT} \cdot N$, $J_{PRT} > 0$.
4) Given $J_{PRT} \in \mathbb{N}$, the index of $\mathbf{x}$ in a partition (PRT) of a $J_{PRT} \cdot N$-point record:
   j_{PRT}.
5) Record times:
   $\{\, t_n \,\}_{n=0}^{N-1}$, where
   $$t_n = n \cdot \Delta t. \tag{1}$$
6) Multitaper bandwidth parameter:
   $$W \in \left(0, \frac{1}{2} \right).$$

Model

1) Generative stochastic process:
   X.
2) Mean signal (SGN) of X:
   $x^{(SGN)}$, with: J_{SGN} normal modes, $J_{SGN} > J_{THR}$; and J_{THR} a threshold (THR) number (under hypothesis).

Parameters

1) The fast Fourier transform (FFT) grid size[1]:
   M_N, where
   $$\log_2 M_N = \lfloor \log_2 N + 1 \rfloor. \tag{2}$$
2) The FFT normalized-frequency vector:
   $\mathbf{f}$, where
   $$\mathbf{f}[m] = \frac{m}{M_N}, \tag{3}$$
   for
   $$m \in \left\{ 0, 1, \cdots, \frac{M_N}{2} \right\}.$$
3) Number of discrete, prolate-spheroidal sequences (DPS) [6]:
   K_N, where
   $$K_N = [\, 2NW - 1 \,]. \tag{4}$$
4) Threshold size, multitaper harmonic F-test:
   α_F, given by
   $$\alpha_F = \frac{1}{N} \vee \min \{\, \alpha \in (0,1) : J_{SGN} > J_{THR} \,\}, \tag{5}$$
   and where $\vee$ is an operator that selects the greater of two values on its either side. Eq. Equation 5 means to say that $(1 - \alpha_F) \cdot 100\%$ is a percentile for which greater than J_{THR} level crossings of the realized F-statistic spectrum occurs. Since α_F is unknown and must be inferred from $\mathbf{x}$, it is necessary to estimate α_F from data. Even though $M_N > N$, only N effective hypothesis tests are conducted, hence the choice of N^{-1} as the nominal α_F-value.

[1] The initialization in Equation 2 is that used in [1].

III. Algorithms

This section relies on quantities that are specified in Section II. All source code discussed in this section is included under the directory, `Harmonic_Anharmonic`. This section makes use of quantities specified in Section II to detail algorithms for multitaper reconstruction of the following quantities and signals.

- The $x^{(SGN)}$ mode frequencies (Section III-A).
- $x^{(SGN)}$ itself (Section III-B).
- The $x^{(SGN)}$ period (Section III-C).

A. Mean-signal Mode Frequencies

Algorithm III-A.1 explains how the multitaper harmonic F-tests over non-negative FFT-frequency are collectively used to obtain an estimate of α_F given $\mathbf{x}$. Starting with the first collection of J_{SGN} identified mode frequencies, Algorithm 2 then can be used to screen the mode-frequency estimates for those at the maxima of spectral peaks and which are resolved at the $2W$-level. In the algorithm, vec is an operator on a finite set that produces a vector whose entries are the set elements - appearing in the same order as in the set originally.

1) Starting Estimates:

Parameters:

1) α_F-estimate given $\mathbf{x}$:
   $\check{\alpha}_F$, making use of Equation 5.

2) $(1-\check{\alpha}_F)$'th quantile of the $F_{2,2K_N-2}$-distribution [2]:
   $F_{2,2K_N-2;1-\check{\alpha}_F}$.

3) Number of F-statistic crossings of $F_{2,2K_N-2;1-\check{\alpha}_F}$, and an estimate of J_{SGN}:
   $\check{J}_{SGN} \in \mathbb{N}_{(J_{THR},30]}$, where the upper bound of 30 prevents large numbers of normal modes being included in the final reconstruction of $x^{(SGN)}$ on $\{t_n\}_{n=0}^{N-1}$.

4) The number of $\check{\alpha}_F$ updates (UPD), in an iterative $\check{\alpha}_F$ estimation that implements Equation 5:
   N_{UPD}.

5) A stopping (STP) Boolean to end a while loop associated with the iterations in the implementation of Equation 5:
   j_{STP}, where

$$j_{STP} = \neg H\left(J_{THR} - \check{J}_{SGN}\right), \tag{6}$$

   and with the terms given as follows.
   - $\neg$ is the negation operator in Boolean algebra.
   - H is the Heaviside function.

Results:

1) Multitaper F-statistic values on the non-negative principal domain:

   $\check{\mathbf{F}} \in \mathbb{R}_{\geq 0}^{\frac{M_N}{2}}$.

2) FFT-indices corresponding to FFT-frequencies nearest to identified mode frequencies of $x^{(SGN)}$:
   $\mathbf{m}_{SGN}$.

[2] Null distribution of the multitaper harmonic F-test, as specified in [7].

Source code: `multitaper_harmonic_F_test.function`.

Initialize: $J_{THR} = 2$; $\check{\alpha}_F = \frac{1}{N}$; $N_{UPD} = 0$; $j_{STP} = 0$; $\mathbf{m}_{SGN}^{\mathrm{T}} = [m]_{m=0}^{\frac{M_N}{2}}$.

Initialize: Reconstruction, $\check{\mathbf{F}}$, of $\mathbf{F}$ (Section D.3 in [1]).

Update: $\mathbf{m}_{SGN} = \mathrm{vec}\left\{ m : \check{\mathbf{F}}[m] > F_{2,2K_N-2;1-\check{\alpha}_F} \right\}$.

Initialize: $\check{J}_{SGN} = \dim\left(\mathbf{m}_{SGN} \right)$.

while $\left(j_{STP} = 0 \text{ **or** } \check{J}_{SGN} < 30 \right)$ **and** $N_{UPD} \leq 10$ **do**

  Update: Either initialize $\check{\alpha}_F$ $\left(N_{UPD} = 0 \right)$ or else increase it by an order of magnitude $\left(N_{UPD} \in \mathbb{N} \right)$:

  $\log_{10}\left(1 - \check{\alpha}_F \right) = -\left\lceil \left| 1 - \log_{10}\left(1 - \frac{1}{N} \right) \right| \right\rceil + N_{UPD}$.

  Update: Be generous how many normal modes to retain - keep $\check{\alpha}_F$ appreciable, $\check{\alpha}_F = 0.05 \vee \check{\alpha}_F$.

  Update: $\mathbf{m}_{SGN}$, $\check{J}_{SGN}$.

  Update: $N_{UPD} = N_{UPD} + 1$.

  if $J_{SGN} > J_{THR}$ **then**
    $j_{STP} = 1$.

  end

end

Algorithm 1: Iterated, data-adaptive inference for $x^{(SGN)}$ normal-mode FFT-frequency indices.

2) Screened Estimates:

This algorithm follows that of Section III-A.1.

Parameters:

1) A $\left(2 \cdot \lfloor M_N W \rfloor \right) \times \check{J}_{SGN}$ matrix, $\mathbf{m}_{NGB}$, of FFT-indices nominal peak-bandwidth neighbourhoods[3], where

$$\mathbf{m}_{NGB}[u, l] = \mathbf{m}_{SGN}[l] - \lfloor M_N W \rfloor + u. \tag{7}$$

Results:

1) Revised versions of $\mathbf{m}_{SGN}$ and $\mathbf{m}_{NGB}$.

2) Vector of $x^{(SGN)}$ mode FFT-frequencies: $\check{\mathbf{f}}_{SGN}$, satisfying the following.

   a) $\check{\mathbf{f}}_{SGN}[l] = \mathbf{f} \circ \mathbf{m}_{SGN}[l]$.

   b) $\check{\mathbf{F}}_{SGN}[l] = \check{\mathbf{F}} \circ \mathbf{m}_{SGN}[l]$.

   c) The p-value of each $\check{\mathbf{F}}_{SGN}$-entry is strictly bounded above by $\check{\alpha}_F$.

   d) Each pair of $\check{\mathbf{f}}_{SGN}$-entries correspond to F-statistic peaks that are resolved at a $2W$-level:

$$\check{\mathbf{f}}_{SGN}[m+1] - \check{\mathbf{f}}_{SGN}[m] > 2W. \tag{8}$$

[3]Refer to [1] for details about choosing the peak bandwidth, where some motivation is provided from a number of references.

Source code: `harmonic_analysis.function`.

Initialize: $\check{\mathbf{F}}$, $\mathbf{m}_{SGN}$, as from the results of Algorithm 1.

Exclude from $\mathbf{m}_{NGB}$ a column if there exists a previous column whose entries collide with those of the presently-considered column.

Initialize: $\check{K}_{SGN} = \check{J}_{SGN}$.

for $l = 2$ **to** $\check{J}_{SGN}$ **by** 1 **do**

  if $\check{\mathbf{f}}_{SGN}[l] - \mathbf{f} \circ \mathbf{m}_{SGN}[l-1] \leq 2W$ **then**

    for $k = 2$ **to** $2 \cdot \lfloor M_N W \rfloor$ **by** 1 **do**

      Initialize:

      $\mathcal{C}_{kl} = \{ \mathbf{m}_{NGB}[k,u] \ : \ u \in \{1, 2, \cdots, \check{K}_{SGN}\} \setminus l \}$.

      Update:

      $\left(\mathbf{m}_{NGB}\left[k, 1 : \left(\check{J}_{SGN} - 1\right)\right]\right)^{\mathrm{T}} = \mathrm{vec}\, \mathcal{C}_{kl}$.

    end

    Initialize:

    $\mathcal{C}_l = \{ \mathbf{m}_{NGB}[u] \ : \ u \in \{1, 2, \cdots, \check{J}_{SGN}\} \setminus l \}$.

    Update: $\mathbf{m}_{SGN}{}^{\mathrm{T}} = \mathrm{vec}\, \mathcal{C}_l$.

    Update: $\check{K}_{SGN} = \check{K}_{SGN} - 1$.

  end

end

Update: $\check{J}_{SGN} = \check{K}_{SGN}$.

Screen frequency estimates further so that each corresponds to the greatest value of an F-statistic peak.

Initialize: $\mathcal{C}_W = \{ 1 : (2 \cdot \lfloor M_N W \rfloor) \}$.

for $l = 1$ **to** $\check{J}_{SGN}$ **by** 1 **do**

  $m_l^* = \underset{j \in \mathcal{C}_W}{\mathrm{argmax}}\, \check{\mathbf{F}} \circ \mathbf{m}_{NGB}[j,l]$.

  Initialize: $\mathcal{C}_W^{(l)} = \{ \mathbf{m}_{NGB}\left[j - \mathbf{m}_{SGN}[l] + m_l^*, l \right] \}_{j \in \mathcal{C}_W}$.

  Update: $\mathbf{m}_{NGB}[1 : |\mathcal{C}_W|, l] = \mathrm{vec}\, \mathcal{C}_W^{(l)}$.

  Update: $\mathbf{m}_{SGN}[l] = \mathbf{m}_{SGN}\left[m_l^* \right]$.

end

Algorithm 2: Updated frequency-estimate FFT-indices, the changes ensuring the frequency estimates are: resolved at the multitaper bandwidth; and at the optimal values of F-statistic peaks.

B. Mean-signal Reconstruction

This algorithm follows that of Section III-A.2, and it yields reconstructions of the $\check{J}_{SGN}$ normal modes of $x^{(SGN)}$.

Result:
A class of $x^{(SGN)}$ normal-mode reconstructions on $\{ t_n \}_{n=0}^{N-1}$: $\mathcal{C}_{SGN}$, where

$$\mathcal{C}_{SGN} = \left\{ \tilde{\mathbf{x}}_l^{(SGN)} \right\}_{l=1}^{\check{J}_{SGN}}. \tag{9}$$

Source code: `basic_mtsa_single_section.function`.

for $l = 1$ **to** $\check{J}_{SGN}$ **by** 1 **do**

  Initialize: $\check{\mathbf{x}}_l^{(SGN)}$ (algorithm of [2], except using as the mode-frequency estimates those obtained from Algorithm 2).

  Update:

  $\check{\mathbf{x}}_l^{(SGN)} = \tilde{\mathbf{x}}_l^{(SGN)}$, with $\tilde{\mathbf{x}}_l^{(SGN)}$ specified as follows.

  - The 1/6'th and 6/6th subvectors of $\check{\mathbf{x}}_l^{(SGN)}$ are replaced by linear-extrapolation reconstructions.
  - Training for the two linear-extrapolation reconstructions discussed in the previous list item:
    - 1/6'th subvector: 2/6'th subvector, first 10 points.
    - 6/6'th subvector: 5/6'th subvector, last 10 points.
  - Extrapolation is performed using linear regression in `R` on each training set of the previous list item and extending both linear trends across the respective replacement gaps.

end

Algorithm 3: Reconstruction of $x^{(SGN)}$ normal modes on $\{ t_n \}_{n=0}^{N-1}$.

C. Period Estimation for the Cosine-series Mean Signal

This algorithm follows that of Section III-A.2. It yields a data-adaptive estimate of the period of $x^{(SGN)}$.

Parameters:

1) A class of FFT-index vectors: $\mathcal{M}_{SGN}$, where
$$\mathcal{M}_{SGN} = \left\{ \mathbf{m}_{SGN}^{(u)} \right\}_{u=1}^{U_{SGN}}, \tag{10}$$
and where the $\mathbf{m}_{SGN}^{(u)}$ are nonoverlapping, contiguous $\mathbf{m}_{SGN}$-subvectors.

2) A class of FFT-frequency vectors: $\mathcal{F}_{SGN}$, where
$$\mathcal{F}_{SGN} = \left\{ \mathbf{f}_{SGN}^{(u)} \right\}_{u=1}^{U_{SGN}}, \tag{11}$$
and where
$$\mathbf{f}_{SGN}^{(u)}[l] = \mathbf{f} \circ \mathbf{m}_{SGN}^{(u)}[l]. \tag{12}$$

3) Augmented vectors, each containing a stack of the first j vectors in $\mathcal{M}_{SGN}$ and $\mathcal{F}_{SGN}$:
$\underline{\mathbf{m}}_{SGN}^{(j)}$ and $\underline{\mathbf{f}}_{SGN}^{(j)}$, respectively.

4) The direction vector for the linear trendline passing through the scatter whose x-coordinates are from $\mathbf{m}_{SGN}^{(u)}$ and whose y-coordinates are from $\mathbf{f}_{SGN}^{(u)}$: $\mathbf{v}_{SGN}^{(u)}$.

5) The direction vector for the linear trendline passing through the scatter whose x-coordinates are from $\underline{\mathbf{m}}_{SGN}^{(u)}$ and whose y-coordinates are from $\underline{\mathbf{f}}_{SGN}^{(u)}$: $\underline{\mathbf{v}}_{SGN}^{(u)}$.

6) A revised integer-separation value between index entries of $\mathbf{m}_{SGN}^{(u)}$: m_u^*.

Result:
Estimate of the fundamental frequency, f_{SGN}, of $x^{(SGN)}$, given: a zero-intercept constraint; and the scatter plot whose x-coordinates are from $\underline{\mathbf{m}}_{SGN}^{(U_{SGN})}$ and whose y-coordinates are from $\underline{\mathbf{f}}_{SGN}^{(U_{SGN})}$: $\check{f}_{SGN}$.

Excel file: `Subject_C5_Harmonic.xlsx`.

Initialize: $K_1 = \dim\left(\mathbf{f}_{SGN}^{(1)}\right)$.

Initialize: $\left(\mathbf{m}_{SGN}^{(1)}\right)^{\mathrm{T}} = [k]_{k=1}^{K_1}$.

Initialize: $\underline{\mathbf{m}}_{SGN}^{(1)}$, $\mathbf{v}_{SGN}^{(1)}$, $\underline{\mathbf{v}}_{SGN}^{(1)}$, $\mathcal{M}_{SGN}$.

for $u = 2$ **to** U_{SGN} **by** 1 **do**
  Initialize: $K_u = \dim\left(\mathbf{f}_{SGN}^{(u)}\right)$.
  Initialize: $\left(\mathbf{m}_{SGN}^{(u)}\right)^{\mathrm{T}} = \mathbf{m}_{SGN}^{(u-1)}[K_{u-1}] + [k]_{k=1}^{K_u}$.
  Initialize: $\underline{\mathbf{m}}_{SGN}^{(u)}$, $\mathbf{v}_{SGN}^{(u)}$, $\underline{\mathbf{v}}_{SGN}^{(u)}$, $\mathcal{M}_{SGN}$.
end

for $u = 2$ **to** U_{SGN} **by** 1 **do**
  if $\exists\, m_u^* : \underline{\mathbf{v}}_{SGN}^{(u-1)} \parallel \mathbf{v}_{SGN}^{(u)}$ **then**
    for $k = u$ **to** U_{SGN} **by** 1 **do**
      Update: $\mathbf{m}_{SGN}^{(k)}[1] = \mathbf{m}_{SGN}^{(k-1)}[K_{k-1}] + m_u^*$.
      for $j = 2$ **to** K_k **by** 1 **do**
        Update: $\mathbf{m}_{SGN}^{(k)}[j] = \mathbf{m}_{SGN}^{(k)}[j-1] + m_u^*$.
      end
      Update: $\underline{\mathbf{m}}_{SGN}^{(k)}$, $\mathbf{v}_{SGN}^{(k)}$, $\underline{\mathbf{v}}_{SGN}^{(k)}$, $\mathcal{M}_{SGN}$.
    end
  end
end

Initialize: $\check{f}_{SGN}$.

Algorithm 4: Given a scatter plot of mode frequency against sequence index, infer the fundamental frequency of $x^{(SGN)}$.

In Algorithm 4, the step size, m_u^* is chosen based on a visual inspection of the scatter plot of mode frequency ($\underline{\mathbf{f}}_{SGN}^{(j)}$) versus sequence index ($\underline{\mathbf{m}}_{SGN}^{(j)}$). Frequency estimates were obtained from `Voltage_Fisher_Statistics.txt`. For the example code, all computations for the period-estimation exercise have been performed using Microsoft Excel [3], using the files,
`Marshall_2023_Period_Estimation_Computations_Section_2.xlsx`
and
`Marshall_2023_Period_Estimation_Computations_Section_20.xlsx`.
In each of those files is found two tables accompanied by respective scatter plots. The left table shows sequence index ($\mathcal{M}_{SGN}$, the `index` column) and identified mode frequency (the `frequency` column). Up until the first boldfaced cell entry of sequence index, the `index` column displays the entries of $\mathbf{m}_{(SGN)}^{(1)}$. Starting at the first boldfaced entry and ending at the cell entry immediately above that which is next boldfaced, the cells display the entries of $\mathbf{m}_{(SGN)}^{(2)}$ - and so forth. To decide on which `index` cell entries were boldfaced, the right table and accompanying right plot were used. The right table and associated right plot have together been obtained using Algorithm 4. The algorithm was implemented as follows.

1) In the left plot, the first 6 points of the scatter exhibit linear correlation (corresponding to $\mathbf{v}_{SGN}^{(1)}$), so there is no need to adjust those points in the right plot (these points correspond to $\mathbf{m}_{(SGN)}^{(1)}$).
2) Hovering the computer cursor over the 7'th point (the user does not know a priori that this point is the 7'th) causes Excel to display the point coordinates. From the 7'th point until the 9'th (these points correspond to $\mathbf{m}_{(SGN)}^{(2)}$), a linear trend (corresponding to $\mathbf{v}_{SGN}^{(2)}$) stands out that differs from both what precedes and what follows in the scatter plot. As such, all scatter points beyond the 6'th were modified so that their time index step was 5 instead of 1 (in both the left and right plots the 6'th scatter point - a blue diamond - is overlaid by an orange disc that itself represents a changepoint). This modification aligns $\mathbf{v}_{(SGN)}^{(2)}$ with $\mathbf{v}_{(SGN)}^{(1)}$, but causes all remaining $\mathbf{v}_{(SGN)}^{(j)}$ not to align as well with $\underline{\mathbf{v}}_{(SGN)}^{(2)}$. Thus, the previous step of modifying $\mathbf{m}_{(SGN)}^{(2)}$ is repeated for the 10'th through 23rd scatter points (these points correspond to $\mathbf{m}_{(SGN)}^{(3)}$). This last step is repeated until all scatter points tightly cluster about the trend line, subject to a low number of changepoints (otherwise, the presence of many changepoints suggests that more than one cosine series superposes to explain $x^{(SGN)}$).

IV. Initialization

All initialization discussed in this section is included under the subdirectory,
`Parameter_Initialization`,
of the source-code directory (i.e., either
`Raw_Series_Plot_Epochs`
or
`Harmonic_Anharmonic`).

A. Considered Quantities and Signals

In addition to the quantities which are specified in Sections II and III, the following are considered in the initialization files.

1) Multitaper spectral-power estimates with M_N-point zero-padding:

   $\hat{\mathbf{s}}_\mu \in \mathbb{R}_{>0}^{\frac{M_N}{2}}$.

2) Signal estimates:
   $\check{\mathbf{x}}^{(SGN)}$, given by

$$\check{\mathbf{x}}^{(SGN)} = \sum_{l=1}^{J_{SGN}} \tilde{\mathbf{x}}_l^{(SGN)}. \tag{13}$$

   In Equation 13, the $\tilde{\mathbf{x}}_l^{(SGN)}$ are the $\mathcal{C}_{SGN}$-elements given in Algorithm 3.
3) Noise reconstructions:
   $\check{\mathbf{x}}^{(NSE)}$, given by

$$\check{\mathbf{x}}^{(NSE)} = \mathbf{x} - \check{\mathbf{x}}^{(SGN)}. \tag{14}$$

In the example code, the time-series parameters are specified as follows (M_N computed in both cases using the specification in Equation 2).

- Full time series:
  $N = 6 \times 10^4$, $M_N = 1.31072 \times 10^5$.
- Section for the 50%-overlap harmonic $x^{(SGN)}$-reconstruction (Algorithm 3):
  $N = 3 \times 10^4$, $M_N = 6.5536 \times 10^4$.

Tables I through VI specify the initialized quantities and Booleans, along with their associated variable names in code files. By adjusting these parameter specifications, the user controls the parameters of basic multitaper power-spectral and harmonic analyses (from controlling the plotting labels to deciding whether to include plots of the spectral-power estimates under a jackknifing estimation regime). File names provided in the six tables and accompanying text refer to files that are situated in subdirectories specified in Section V-A.

B. Boolean Parameters

The titles of Tables I and II specify the files under the `Boolean_Parameters` directory where are initialized the parameters of mention. Table I includes parameters that specify whether to perform certain spectral analyses, while Table II parameters are used to decide whether to produce output of certain spectral signals that characterize the X finite-dimensional distribution

TABLE I
BOOLEANS.R

Question	Variable
$J_{PRT} \in (0,1)$?	`truncation_bool.par`
Compute $\hat{\mathbf{s}}_\mu$?	`spectral_power.bool`
Harmonic analysis for $\mathbf{x}$?	`ha_bool.par`
Power spectral analysis for $\check{\mathbf{x}}^{(NSE)}$?	`residual_eigencoeffs.bool`

For Table II below regarding spectral analysis, some of the output Boolean variables have been set to `FALSE` in the example code, though the user may wish to assign `TRUE` to those if only a handful of section series (i.e., $J_{PRT} \in \mathbb{N}_{>1}$) is to be considered and the user would like to perform analysis specifically on the associated file contents.

TABLE II
SPECTRAL_ANALYSIS_BOOLEANS.R

Question	Variable
Jackknifed spectral-power estimation?	`jk_spectral_power.bool`
Output `Voltage_Spectral_Power_Estimates.txt`, `Residual_Voltage_Spectral_Power_Estimates.txt`, and `Voltage_Fisher_Statistics.txt`?	`output.spectral_power`
Output `Voltage_Jackknifed_Spectral_Power_Estimates.txt` and `Residual_Voltage_Jackknifed_Spectral_Power_Estimates.txt`?	`output.jk_spectral_power`

Algorithm 5 below presents a logic algorithm (itself contained in the `Parameter_Initialization.R` file under `Parameter_Initialization`) that ensures basic requirements are met if the user only initializes a few Boolean variables that depend on others in order to provide the required output (e.g., if the user desires to obtain a data file that includes the power spectrum and associated FFT frequencies, then setting to TRUE the output Boolean necessitates that the Boolean for a spectral analysis be set to TRUE as well).

Source code: `basic_mtsa_single_section.function`.

if

$(\neg \text{output.jk\_spectral\_power} = 0) \cup$

$(\neg \text{jk\_spectral\_power.bool} = 0) \cup$

$(\neg \text{output.jk\_spectral\_power} = 0)$

then

- $\neg \text{spectral\_power.bool} = 0.$
- $\neg \text{spectral\_power\_bool.par} = 0.$
- $\neg \text{ha\_bool.par} = 0.$

end

if

$(\neg \text{residual\_eigencoeffs.bool} = 0) \cup$

$(\neg \text{output.harmonic\_Fisher\_statistics} = 0)$

then

$\neg \text{ha\_bool.par} = 0.$

end

Algorithm 5: Logic conditional on what the user requires of the spectral and harmonic analyses.

C. Graphical Parameters

The title of Table III below specifies the file under the `Output_Graphical_Parameters` directory where are initialized the parameters of mention. By adjusting the `plotting.time_scale` and `first_time_units.string` parameters in Table III, the user changes the units in both the `Section_Voltage_Series_Raw.pdf` and `Voltage_Series_Raw_Sections.pdf` files, but not in the `Harmonic_Reconstruction_Series.pdf` and `Harmonic_Residuals_Series.pdf` files (in the latter two files, the units of the base time remain unchanged from those of the corresponding time axis - all in raw data units). This formatting discrepancy in the harmonic-reconstruction files should be addressed in a future code upgrade, though it is not a significant issue for intents and purposes; the plots for which the specific formatting has been applied are the primary ones for which this formatting is beneficial where is concerned assessment of the survey strata.

TABLE III
OUTPUT_GRAPHICAL_PARAMETERS.R

Variable / Description
`plotting.time_scale`, `first_time_units.string` The first parameter is a conversion scale factor from units of the time axis to units of the base time. In the example code, the time axis is in seconds (`time_units.string=''seconds''`), but the base time is on the order of 10^3 s. To write the base time in minutes, the required code is `plotting.time_scale=1/60` `first_time_units.string=''minutes''`.
`num.plotting_points` Maximum number of points of $\mathbf{x}$ for ease of visualization in the plotting (or a downsampled version of $\mathbf{x}$, if N exceeds this limit).
Time-domain parameters • Time – `measured_abscissa.string`: The time quantity of the time series (e.g., "Time" or "Space"). – `time_units.string`: Time units (e.g., "seconds" or "metres"). – `abbreviated_time_units.string`: Abbreviated time units (e.g., s or m). • State – `measure_quantity.string`: The state quantity of the time series (e.g., "Voltage" or "Flux"). – `measure_quantity_string.lower_case`: Same as `measure_quantity.string`, except with lowercase initial letter. – `measure_units.string`: State units (e.g., "volts" or "particles per squared metre"). Frequency-domain parameters • `frequency_units.string`: Frequency units (e.g., "cycles per second" or "cycles per metre"). • `abbreviated_frequency_units.string`: Abbreviated frequency units (e.g., cps or cpm).

D. Time-series Parameters

1) General Parameters:

The titles of Tables IV and V specify the files under the `Time_Series_Parameters` directory where are initialized the parameters of mention.

In Table IV below,

$$\text{N.par} = 6 \times 10^4$$

in the example code because N this large ensures high-resolution frequency-domain statistics in a multitaper spectral analysis [4]. From Table V,

$$\text{F\_test.threshold} = \text{NA},$$

which itself enforces the relevant functions in the `General_Headers` directory files to make the initialization of $\check{\alpha}_F$ such that $J_{THR} = 2$ [5].

TABLE IV
TIME_SERIES_PARAMETERS.R

Quantity	Variable
N	`N.par`

TABLE V
MULTITAPER_PARAMETERS.R

Quantity	Variable
NW	`NW.par`
J_{THR}	`F_test.threshold`

2) Specific Parameters:

The title of Table VI specifies the file under the `Application_Specific_Parameters` directory where is initialized the sole parameter of mention. The parameter is a concatenation of j_{PRT}-values. That is, the example code has `selected_time_window.indices<-c(2,11,20,32)`, As a special example, the user could use the initialization, `selected_time_window.indices<-c(20)`, in order to analyze the 20'th section alone - which is tantamount to the initialization,

$$j_{PRT} = 20.$$

TABLE VI
APPLICATION_DATA_PARAMETERS.R

Quantity	Variable
A collection of j_{PRT} values.	`selected_time_window.indices`

[4]The motivation behind this present statement arises from record-size considerations in a probabilistic analysis from Chapters 4 and 5 of [1].

[5]This result is achieved using Algorithm 1.

V. Code Instructions

This section contains mathematical symbols from Sections II and IV-A. It provides the following.

1) Navigation through the directory contents.
2) Instructs to run the program (e.g., sourcing `Run_Code.R`).
3) Ways to adjust the code in order to make modifications to the example analysis (e.g., applying the code to a different dataset).

A. Starting Considerations

1) Download the following two Excel spreadsheets.

   - `Marshall_2023_Period_Estimation_Computations_Section_2.xlsx`.
   - `Marshall_2023_Period_Estimation_Computations_Section_20.xlsx`.

2) Download three code directories into the working directory of choice (ensure that you situate these two directories in a location near your `Documents` directory [6] details of the directory contents provided below).
   a) **General_Headers**
   All header files whose functions are used in execution of the algorithms of Section III. In addition, the directory contains functions for: multitaper spectral analysis; time-series analysis; file input and output; plotting; general statistical methods; and other analysis methods. The details of the functions will be further provided in the near future, with manuscripts in writing that apply these functions in example analyses.
   b) Two directories containing main files.
   - **Raw_Series_Plot_Epochs**
   Contains the following directories.
     i) **Parameter_Initialization**
     ii) **Setup**
       – Run_Code
         * Run_Code.R
         Sourcing this code executes the entire `R` program.
     iii) **Data**
       – Time_Series
         * C5_electrode_1.txt
         A single time-series data text file, containing a single, untitled column of ordinate values.
       – Parameters
         * Patient_C5.txt
         A text file for the abscissa associated with the time series in `C5_electrode_1.txt` above. The two lines of this text files are specified as follows.
         i) Base time, in units of `time_units_string.par` from Table III. This is the time to which add the t_n [7].
         ii) The sampling frequency, in units of `frequency_units.string` from Table III.

[6] Otherwise, on a Windows operating system - perhaps also on a Macintosh operating system - the directory naming might be too long and incur an error.
[7] Recall from Equation 1 how the t_n satisfy $t_0 = 0$

iv) **1_Section**
The title holds for the case that `selected_time_window.index=1`, but the assigned value of `selected_time_window.index` replaces the "1" at the start of this directory title if it differs from unity [8].

- 1_Preliminary_Analysis
  - Section_Voltage_Series_Raw.pdf
    The raw time series, $\mathbf{x}$.
  - Voltage_Series_Raw_Sections.pdf
    The same plot as in Section_Voltage_Series_Raw.pdf, but now containing multiple pages: one per section of size N from the global record ($\mathbf{x}$ in this case refers to a section of the global record, rather than the global record itself).
  - Total_Computation_Time.txt
    Provides the total computation time during code execution.

- **Harmonic_Anharmonic**
  Same as the Raw_Series_Plot_Epochs directory, except that: a) it contains two files worth considering; and b) its 1_Preliminary_Analysis/2_Section subdirectory contains different content from the kind contained in the 1_Section/1_Preliminary_Analysis, subdirectory of Raw_Series_Plot_Epochs, as follows.
  - Output_File.txt
    Contains output such as: the timing of certain computations (e.g., spectral-power estimation); number of $\mathbf{x}$ data points used for analysis; FFT grid size (i.e., M_N); and sizes of the series used in harmonic reconstruction of $x^{(SGN)}$.
  - Section_Voltage_Series_Raw.pdf
    The same plot as for the second page in Voltage_Series_Raw_Sections.pdf (corresponding to the second section of the global record), but now only including the page whose number is the assigned value of the entry of selected_time_window.indices. that specifies the second $\left(6 \times 10^4\right)$-point section of the global record.
  - **1_Section/Multitaper_Spectral_Power**
    - Voltage_Jackknifed_Spectral_Power_Estimates.txt
      Jackknifed estimates of X power-estimator mean (average) and associated confidence bounds (LB, UB): point-by-point over the non-negative part of the principal domain (frequency).
    - Voltage_Jackknife_Spectral_Power_Estimates.pdf
      Contents of the above text file.
      · average: Black curve.
      · LB, UB: Bounds of the grey-filled region.
    - Residual_Voltage_Jackknife_Spectral_Power_Estimates.txt
      Same as Voltage_Jackknife_Spectral_Power_Estimates.txt, except now for $X^{(NSE)}$ instead of X. Here, the multitaper power estimator has the same form as previously for the X case, but the eigencoefficient spectra have been reconstructed by subtracting from the $\mathbf{x}$ eigencoefficient spectra multitaper reconstructions of the associated $x^{(SGN)}$ eigencoefficient spectra [9].
    - Residual_Voltage_Jackknife_Spectral_Power_Estimates.pdf
      Same as Voltage_Jackknife_Spectral_Power_Estimates.pdf, except now for $X^{(NSE)}$ instead of X.
    - Voltage_Spectral_Power_Estimates.txt
      Same as Voltage_Jackknife_Spectral_Power_Estimates.pdf, except now with: no confidence-bound estimates included; and the mean estimates having been replaced by the $\check{\mathrm{s}}_\mu$-estimates (spectral_power).
    - Voltage_Spectral_Power_Estimates.pdf
      Plot of the contents of Voltage_Spectral_Power_Estimates.txt
    - Residual_Voltage_Spectral_Power_Estimates.txt
      Same as Voltage_Spectral_Power_Estimates.txt, except now for $X^{(NSE)}$ instead of X.
    - Residual_Voltage_Spectral_Power_Estimates.pdf
      Same as

[8] For the example code, it is best to leave unchanged the initialization of this parameter because it is only of use in other analyses for certain manuscripts in writing.

[9] Refer to Appendix D.3 in [1] for more details.

Voltage_Spectral_Power_Estimates.pdf,
except now for $X^{(NSE)}$ instead of X.

* Voltage_Fisher_Statistics.txt
The spectrum of statistically-significant multitaper harmonic F-statistic values (SNR, since the F-statistic is a metric associated with the signal-to-noise ratio [10]). Accompanying the significant F-statistic values are the percentiles (percentile) of the null $F_{2,2K_N-2}$ null distribution.
* F_spectrum.pdf
Plot of the contents of Voltage_Fisher_Statistics.txt and for the F-statistic spectrum covering the rest of the non-negative principal domain. Accompanying are: a single, horizontal line, drawn at the $\check{\alpha}_F$; and a number of vertical lines, drawn at the frequencies corresponding to level normal-mode detections from those multitaper harmonic F-tests which have been conducted point-by-point over the principal domain.
* Harmonic_Reconstruction_Series.pdf
The raw time series, $\mathbf{x}$, overlaid by the harmonic-mean reconstruction, $\check{\mathbf{x}}^{(SGN)}$.
* Harmonic_Residuals_Series.pdf
The $\check{\mathbf{x}}^{(NSE)}$-reconstruction.
* Means_Ends_Reconstruction/Interpolation
Contains results from a code useful for some multitaper analyses, but not the one considered in the example code. This directory is of interest for the analyses of certain manuscripts in writing, where hard-replacement interpolation is important (e.g., it is useful for the interpolation of gaps in the practical examples discussed in [1]). Refer to Section V-D for further details regarding the contents of this directory.

3) Delete the 1_Preliminary_Analysis subdirectory, or else keep this template directory in a separate location for your own reference.
4) Replace:
   - C5_electrode_1.txt with the data file of interest; and
   - Patient_C5.txt with the parameter file of interest.

   In both cases, preserve the forms of the file structure, as specified above in 1 b iii). In addition, keep the names of these two files distinct.
5) One can change the Harmonic_Anharmonic title of the application code to the name of choice, so long as one adjusts the following lines accordingly.
   a) In Run_Code.R,
   adjust the initializing string in the line,
   code.string<<-''Harmonic_Anharmonic''
   b) Change the title of the Harmonic_Anharmonic.R code file from Harmonic_Anharmonic to the contents of the new string used for the above code.string initialization.
6) In the file,
Application_Data_Parameters.R,
under the Application_Specific_Parameters subdirectory, adjust the following line depending on the choice of j_{PRT} (here in these directory-content explanations, $j_{PRT} = 2$).
selected_time_window.indices<<-c(2)
The left table column displays which of the epochs in Voltage_Series_Raw_Sections.pdf have been selected (in the example code,

$$J_{PRT} = 40,$$

while

$$j_{PRT} \in \{2, 11, 20, 32\}).$$

The right table column provides the names of the epochs for which the section time series indexed in the left column have been deemed representative (this is illustrated in Table VII).

TABLE VII
INITIALIZATION FOR SELECTED_TIME_WINDOW.INDEX: AN ILLUSTRATIVE EXAMPLE

Section No.	Epoch type
2	Alpha periodicity
11	Ictal wavefront
20	Gamma periodicity
32	Ictal discharges

[10]Refer to [1] for details of said relationship.

B. Harmonic Analysis: `Harmonic_Anharmonic`

If the size of the global record in the `C5_electrode_1.txt` data file is on the order of 10^4 or smaller, then it may be desirable not to section the times series prior to harmonic reconstruction of $x^{(SGN)}$ [11]. To set

$$J_{PRT} = 1$$

and make use of all of the global record (i.e., $\mathbf{x}$ comprises all data points in `C5_electrode_1.txt`), use the initialization of Table VIII. Each row of the first column of the table contains the identifier number of the table in Section IV that specifies the parameter whose initialization is provided in the corresponding right column entry.

TABLE VIII
INITIALIZATION FOR FULL-RECORD ANALYSIS

Table No.	Initialization
I	`truncation_bool.par=FALSE`
IV	`N.par=NA`
VI	`time_window_index.par=1`

C. Runtime Considerations: `Harmonic_Anharmonic`

Table VIII presents output (second column) from the `Output.txt` files under those directories which are associated with the analyses performed on sections of $\mathbf{x}$ whose indices are specified in the first column. That is: the first column lists sections that have been identified using the pages from the `Voltage_Series_Raw_Sections.pdf` file. Evident from the third column: with runtimes ranging 4-5min, restricting the number of identified normal modes to 30 has the effect of leveling the runtime across the epoch stratum. Most compute time is required for the multitaper computations, evenly separated between: the full-record harmonic analysis on $\mathbf{x}$; and the harmonic reconstruction of $x^{(SGN)}$.

TABLE IX
RUNTIMES FOR THE EXAMPLE CODE

Section No.	Seconds	Minutes
2	272.13	4.5
11	304.26	5.1
20	272.71	4.5
32	251.32	4.2

[11] Recall Algorithm 3.

D. Interpolation: Means_Ends_Reconstruction/Interpolation

The Interpolation directory is redundant for the reconstructions considered in the multitaper power-spectral and harmonic analyses of the example code. The code has been recycled from that used to interpolate gaps used in a hard-replacement scheme of [1] [12]. In this case, the gap comprises the central 10^4 points of an $x^{(SGN)}$-reconstruction - which would have been useful only for the case that only the first and third of the overlapping sections had been considered separately for analysis. Moreover, a $\left(1.2 \times 10^5\right)$-point record is under analysis, whereas the example code deals only with a $\left(6 \times 10^4\right)$-point record.

- Neighbours.txt
  Training for harmonic reconstruction of $x^{(SGN)}$ by extrapolation given $\mathbf{x}$ based on three $\left(6 \times 10^4\right)$-point section series with 50% overlap (i.e., $N = 1.2 \times 10^5$).
  gap_index=1: First gap introduced for hard replacement.
  The prior_time through post table columns include the time indices and voltages used in training for an update of the section-overlap harmonic-mean reconstruction in the gap between time indices 5.5×10^4 and 6.5×10^4. The first and second halves of this updated reconstruction are retained for the first and third $\left(6 \times 10^4\right)$-point section reconstructions, respectively. The left training section for gap interpolation (prior_time for time index and prior for voltage) spans the time indices 2.5×10^4 through 5.5×10^4. The first of these two bounding indices occurs one quarter into the 1/6'th section of the first $\left(6 \times 10^4\right)$-point section. As for the right training section (post_time for time index and post for voltage), the training section has been chosen adjacent to and right of gap, using the same number of training points as in the left training section.
  gap_index=2: Second gap introduced for hard replacement.
  The same setup as for the first gap, but now between time indices 1.15×10^5 and 1.25×10^5 (i.e. at the right end of the third $\left(6 \times 10^4\right)$-point section).
- Voltage_Series_Pilot_Interpolation.txt
  Reconstructed values in the gap specified in Neighbours.txt. Here, the time_index and gap_index table headings have the same interpretation as they do in Neighbours.txt.

[12] Something that is of value in a manuscript in writing that involves novel bandpass filters in neuroscience application.

ACRONYMS

DPS discrete, prolate-spheroidal sequences. 2

FFT fast Fourier transform. 2–6, 8, 12

PRT partition. 2

SGN signal. 2

STP stopping. 3

THR threshold. 2

UPD updates. 3

REFERENCES

[1] F. A. Marshall. *Advances in the detection and characterization of non-stationary processes: An application to riometers*. PhD thesis, Queen's University, Kingston, Ontario, 2020.

[2] F. A. Marshall, D. J. Thomson, G. Takahara, and R. A. Fiori. A multitaper model for quiet voltage in relative ionospheric opacity meters. In *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2019.

[3] Microsoft Corporation. Microsoft Excel.

[4] E. D. Schlafly and M. A. Kramer. Human seizure MEA, 2022. Boston College, Accessed: 2022, Last updated 2022-05-25.

[5] Emily D Schlafly, François A Marshall, Edward M Merricks, Uri T Eden, Sydney S Cash, Catherine A Schevon, and Mark A Kramer. Multiple sources of fast traveling waves during human seizures: resolving a controversy. *Journal of Neuroscience*, 42(36):6966–6982, 2022.

[6] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty IV: The discrete case. *Bell System Technical Journal*, 57(5):1371–1430, 1978.

[7] D. J. Thomson. Quadratic-inverse spectrum estimates: Applications to palaeoclimatology. *Phil. Trans. R. Soc. Lond. A*, 332(1627):539–597, 1990.