A NOTE ON MEASURES OF SINGLE TIMESERIES ACTIVITY IN RESTING-STATE FMRI STUDIES

XI-NIAN ZUO

1. The fundamental mathematics

In statistics, signal processing and mathematical finance, a time series is a sequence of data points, measured typically at successive times spaced at uniform time intervals. Here, I only focus on the discrete time series $x_i = x(t_i), i = 1, \cdots, N$. The motivation of writing this methodology note is to clarify the relationship between the measures to quantify activity of an observed timeseries. Particularly, focus one several measures used in recently emerged field of resting-state fMRI studies ([Biswal(1995), Biswal(2010)]). Although all these measures have been extensively developed in EEG or other timeseries studies ([Le Van Quyen(2007)]), many researchers focusing on various applications in or new comers to this new field still need to know existed measures and the relationships among these measures ([Garrett(2010)]).

1.1. **Temporal domain.** There are several basic statistics for timeseries x in below. The mean describes the central location of the data, and the standard deviation describes the spread. The standard deviation remains the most common measure of statistical dispersion, measuring how widely spread the values in a data set are. In mathematics, the root mean square (abbreviated RMS or rms), also known as the quadratic mean, is a statistical measure of the magnitude of a varying quantity.

Mean:

$$x_{\text{mean}} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{1}$$

Standard Deviation:

$$x_{\text{std}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}. \tag{2}$$

Root Mean Square:

$$x_{\text{rms}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}. \tag{3}$$

Theorem 1.1 (The relationship of mean, std and rms)**.**

$$x^2_{\text{rms}} = x^2_{\text{mean}} + x^2_{\text{std}}. \tag{4}$$

Proof. Let's start from the rms,

$$\begin{aligned}
x^2_{\text{rms}} &= \frac{\left[\sum\limits_{i=1}^{N}(x_i - x_{\text{mean}} + x_{\text{mean}})^2\right]}{N} \\
&= \frac{\sum\limits_{i=1}^{N}\left[(x_i - x_{\text{mean}})^2 + 2x_{\text{mean}}(x_i - x_{\text{mean}}) + x^2_{\text{mean}}\right]}{N} \\
&= \frac{\sum\limits_{i=1}^{N}(x_i - x_{\text{mean}})^2 + 2x_{\text{mean}}\sum\limits_{i=1}^{N}(x_i - x_{\text{mean}}) + \sum\limits_{i=1}^{N}x^2_{\text{mean}}}{N} \\
&= \frac{\sum_{i=1}^{N}(x_i - x_{\text{mean}})^2}{N} + 0 + x^2_{\text{mean}} = x^2_{\text{std}} + x^2_{\text{mean}}
\end{aligned}$$

□

Remark. The first thing is, the rms is the same as std if the time series has zero mean. The second is If the sample version of std is used then the relationship comes as

$$x^2_{\text{rms}} = x^2_{\text{mean}} + (1 - 1/N)x^2_{\text{std}}.$$

1.2. **Frequency domain.** In mathematics, the discrete Fourier transform (DFT) is one of the specific forms of Fourier analysis. It transforms one function (here is the time series as the function of time) into another, which is called the frequency domain representation, or simply the DFT, of the original function. Since FFT algorithms are so commonly employed to compute the DFT, the two terms are often used interchangeably in colloquial settings, although there is a clear distinction: "DFT" refers to a mathematical transformation, regardless of how it is computed, while "FFT" refers to any one of several efficient algorithms for the DFT. This distinction is further blurred.

Definition 1.1 (DFT and IDFT)**.** The time series $x_i = x(t_i), i = 1, 2, \cdots, N$ with N sample points in time domain is transformed into a series of N numbers $X_k = X(f_k), k = 1, 2, \cdots, N$ by the DFT according to the formula:

$$X_k = \sum_{i=1}^{N} x_i e^{\frac{-2\pi j}{N}ki}, k = 1, 2, \cdots, N.$$

The inverse discrete Fourier transform (IDFT) is given by

$$x_i = \frac{1}{N}\sum_{k=1}^{N} X_k e^{\frac{2\pi j}{N}ik}, i = 1, 2, \cdots, N.$$

Theorem 1.2 (Parseval's theorem)**.**

$$\sum_{i=1}^{N} |x_i|^2 = \frac{1}{N} \sum_{k=1}^{N} |X_k|^2.$$

Remark. (1). Here x_i are real numbers, then the DFT obeys the symmetry:

$$X_k = X^*_{N-k},$$

where the star denotes complex conjugation. Therefore, the DFT output for real inputs is half redundant, and one obtains the complete information by only looking at roughly half of the outputs. In this case, the "DC" element X_0 is purely real, and for even N the "Nyquist" element $X_{N/2}$ is also real, so there are exactly N non-redundant real numbers in the first half + Nyquist element of the complex output X. (2). All applications of the DFT depend crucially on the availability of a fast algorithm to compute discrete Fourier transforms and their inverses, a fast Fourier transform (FFT). (3). An FFT computes the DFT and produces exactly the same result as evaluating the DFT definition directly; the only difference is that an FFT is much faster. (4). In right side of above equation, there is slight difference when FFT is employed for the fast computation, N could be the actual number of padding, for example is the smallest power of two that is greater than or equal to N.

Definition 1.2 (Power Spectral Density and Power Spectrum)**.** The power spectral density, PSD, describes how the power (or variance) of a time series is distributed with frequency. Mathematically, PSD is the squared modulus of the Fourier transform of the time series, scaled by a proper constant term. The power spectrum, PS, gives a plot of the portion of its power (energy per unit time) falling within given frequency bins. For a given time series x_i, X_k is its DFT, then $\frac{1}{N}|X_k|^2$ is the power spectral density and the power spectrum is

$$\mathrm{PS}(n_1, n_2) = \frac{1}{N} \sum_{k=n_1}^{n_2} |X_k|^2.$$

Denote $\mathrm{PS}(1, N)$=PS. Also usually call $\frac{1}{N}|X_k|$ amplitude spectrum or spectral amplitude, which can be regarded as a L^1-norm PSD.

From the Parseval's theorem, we can get

Theorem 1.3.

$$x_{\mathrm{rms}} = \sqrt{\frac{\mathrm{PS}}{N}}.$$

2. RSFA, LSFA, ALFF and fALFF

In resting state fMRI studies, there are several temporal or frequency measures for analyzing the activities of BOLD time series. Of note, the FFT of a timeseries is still a somehow timeseries in frequency domain, thus RMS or SD can be used to measure activity of the original timeseries. In resting-state fMRI field, Biswal et al. ([Biswal(1995)]) is the first to notice this and use RMS of a timeseries' PSD to quantify the overall gray or white matter activity Mathematically, it is important to clarify the relationships between them. Although there is a frequency-wise method using PSD ([Duff(2008), McAvoy(2008)]). In this section, I only discuss four main measures which summarize the temporal or frequency characteristics of BOLD signals. They are the resting state physiological fluctuation amplitude (RSFA [Kannurpatti(2008)]), low-frequency spectral amplitude (LFSA [Biswal(2007)]), amplitude of low-frequency fluctuations (ALFF [Zang(2007)]) and fractional amplitude of low-frequency fluctuations (fALFF [Zou(2008)]). Of particular importance, due the relationship among these measures, their test-retest reliability should be the similar quantity ([Zuo(2010)]).

Definition 2.1 (RSFA)**.** Given a BOLD time series x, this measure is exactly the same as the standard deviation. Specially, it equals the root mean square if the time series is de-meaned firstly. Therefore, for a de-meaned time series x

$$x_{\text{RSFA}} = x_{\text{std}} = x_{\text{rms}} = \sqrt{\frac{\text{PS}}{N}}.$$

Definition 2.2 (LSFA)**.** With a time series x, its amplitude spectrum at a special frequency $f(k_0)$ is defined as the low-frequency spectral amplitude, i.e.

$$x_{\text{LSFA}}(k_0) = \frac{1}{N}|X_{k_0}|.$$

Definition 2.3 (ALFF)**.** The amplitude of low-frequency fluctuations sums the amplitude spectra within a specific low-frequency band $[f(n_1), f(n_2)]$, i.e.

$$x_{\text{ALFF}}(n_1, n_2) = \frac{1}{\sqrt{N}} \sum_{k=n_1}^{n_2} |X_k|.$$

Definition 2.4 (fALFF)**.** The fractional ALFF is the proportion of low-frequency amplitude spectra in the whole amplitude spectra, i.e.

$$x_{\text{fALFF}}(n_1, n_2) = \frac{x_{\text{ALFF}}(n_1, n_2)}{x_{\text{ALFF}}(1, N)}.$$

Based on the above definitions, as the measures in frequency domain, the Parseval's theorem is the bridge between the time domain and frequency domain. LSFA, ALFF and fALFF are all derived from the L^1-norm PSD. As for L^1-norm PSD, the Parseval's theorem becomes following theorem.

Theorem 2.1 (L^1-norm Parseval's theorem)**.**

$$x_{\text{rms}} = \frac{1}{N}\|\text{PSD}^{L^1}\|.$$

3. Computational and practical issues

There are several software packages to carry out the power spectrum density analysis.

- MATLAB is the most powerful tool to calculate PSD of a timeseries. It includes many functions to compute PSD. For examples, fft, pwelch, pmtm, etc. Another three packages are AFNI, FSL and REST.
- There is a plugin (Power Spectrum) embedded in AFNI's GUI. Some parameters can be modified to optimal PSD computation: Hamming Taper Percent and nFFT. However, it is not available when we want batch processing for large dataset. Of note, AFNI always removes the linear trend in timeseries first and then calculates the PSD.
- FSL has a command referred as 'fslpspec' for PSD computation. Its advantage is available for batch processing. But, there is no optional parameters for FFT calculation. It is based on the classic periodogram method for estimation of power spectrum. There is no linear detrending in fslpspec and usually detrending is suggested before power spectrum estimation. This has been integrated in a set of *nix scripts to calculate both ALFF and fALFF. (see '1000 Functional Connectomes Project': *www.nitrc.org/projects/fcon_1000*)
- REST (*restfmri.net*) is very similar to AFNI's plugin. But it is built in MATLAB, therefore you can call rich functions of PSD estimation to improve your research. Of particular easy-to-use, the developers provide a GUI of REST (DPARSF: [Yan(2010)]).

As for classic periodogram approach, I wrote MATLAB codes to support that computation PSD can be equivalent among the three tools (AFNI, FSL and MATLAB). I chose two BOLD timeseries from two voxels within human resting brain and calculated their PSD using the three softwares. Appendix demonstrates the equivalence among results.

References

Biswal(1995). Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med. 1995; 34(4):537-41.

Biswal(2010). Biswal BB, Mennes M, Zuo XN, Gohel S, Kelly C, Smith SM, Beckmann CF, Adelstein JS, Buckner RL, Colcombe S, Dogonowski AM, Ernst M, Fair D, Hampson M, Hoptman MJ, Hyde JS, Kiviniemi VJ, Kotter R, Li SJ, Lin CP, Lowe MJ, Mackay C, Madden DJ, Madsen KH, Margulies DS, Mayberg HS, McMahon K, Monk CS, Mostofsky SH, Nagel BJ, Pekar JJ, Peltier SJ, Petersen SE, Riedl V, Rombouts SA, Rypma B, Schlaggar BL, Schmidt S, Seidler RD, Siegle GJ, Sorg C, Teng GJ, Veijola J, Villringer A, Walter M, Wang L, Weng XC, Whitfield-Gabrieli S, Williamson P, Windischberger C, Zang YF, Zhang HY, Castellanos FX, Milham MP. Toward discovery science of human brain function. Proc Natl Acad Sci U S A. 2010; 107(10):4734-9.

Le Van Quyen(2007). Le Van Quyen M, Bragin A. Analysis of dynamic brain oscillations: methodological advances. Trends Neurosci. 2007;30(7):365-73.

Duff(2008). Duff EP, Johnston LA, Xiong J, Fox PT, Mareels I, Egan GF. The power of spectral density analysis for mapping endogenous BOLD signal fluctuations. Hum Brain Mapp. 2008; 29(7):778-90.

McAvoy(2008). McAvoy M, Larson-Prior L, Nolan TS, Vaishnavi SN, Raichle ME, d'Avossa G. Resting States Affect Spontaneous BOLD Oscillations in Sensory and Paralimbic Cortex.J Neurophysiol. 2008;100(2):922-31.

Zuo(2010). Zuo XN, Di Martino A, Kelly C, Shehzad ZE, Gee DG, Klein DF, Castellanos FX, Biswal BB, Milham MP. The oscillating brain: complex and reliable. Neuroimage. 2010; 47(4):1579-89.

Zou(2008). Zou QH, Zhu CZ, Yang Y, Zuo XN, Long XY, Cao QJ, Wang YF, Zang YF. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. J Neurosci Methods. 2008; 172(1):137-41.

Zang(2007). Zang YF, He Y, Zhu CZ, Cao QJ, Sui MQ, Liang M, Tian LX, Jiang TZ, Wang YF. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain Dev. 2007; 29(2):83-91.

Kannurpatti(2008). Kannurpatti SS, Biswal BB. Detection and scaling of task-induced fMRI-BOLD response using resting state fluctuations. Neuroimage. 2008; 40(4):1567-74.

Biswal(2007). Biswal BB, Kannurpatti SS, Rypma B. Hemodynamic scaling of fMRI-BOLD signal: validation of low-frequency spectral amplitude as a scalability factor. Magn Reson Imaging. 2007; 25(10):1358-69.

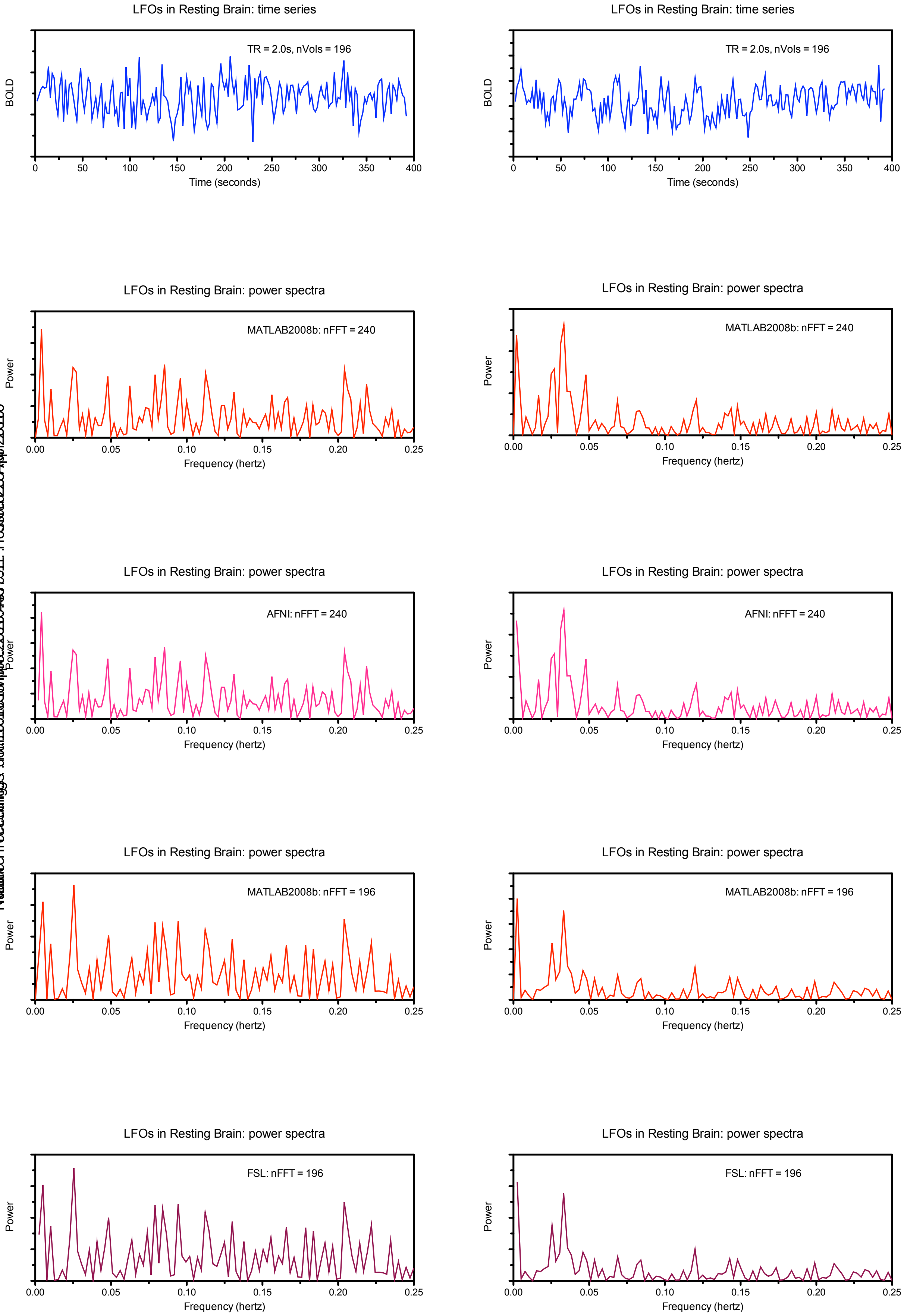
Garrett(2010). Garrett DD, Kovacevic N, McIntosh AR, Grady CL. Blood oxygen level-dependent signal variability is more than just noise. J Neurosci. 2010; 30(14):4914-21.

Yan(2010). Yan CG and Zang YF. DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. Front. Syst. Neurosci. 4:13. doi:10.3389/fnsys.2010.00013

Institute of Pediatric Neuroscience, NYU Child Study Center, 215 Lexington Ave., 14th Floor, New York, NY 10016

E-mail address: xinian.zuo@nyumc.org

URL: http://www.AboutOurKids.org


LFOs in Resting Brain: time series
TR = 2.0s, nVols = 196
BOLD
Time (seconds)
LFOs in Resting Brain: time series
TR = 2.0s, nVols = 196
BOLD
Time (seconds)
LFOs in Resting Brain: power spectra
MATLAB2008b: nFFT = 240
Power
Frequency (hertz)
LFOs in Resting Brain: power spectra
MATLAB2008b: nFFT = 240
Power
Frequency (hertz)
LFOs in Resting Brain: power spectra
AFNI: nFFT = 240
Power
Frequency (hertz)
LFOs in Resting Brain: power spectra
AFNI: nFFT = 240
Power
Frequency (hertz)
LFOs in Resting Brain: power spectra
MATLAB2008b: nFFT = 196
Power
Frequency (hertz)
LFOs in Resting Brain: power spectra
MATLAB2008b: nFFT = 196
Power
Frequency (hertz)
LFOs in Resting Brain: power spectra
FSL: nFFT = 196
Power
Frequency (hertz)
LFOs in Resting Brain: power spectra
FSL: nFFT = 196
Power
Frequency (hertz)