Neural Signal Processing and Spectral Analysis

Partha P Mitra

Cold Spring Harbor Laboratory

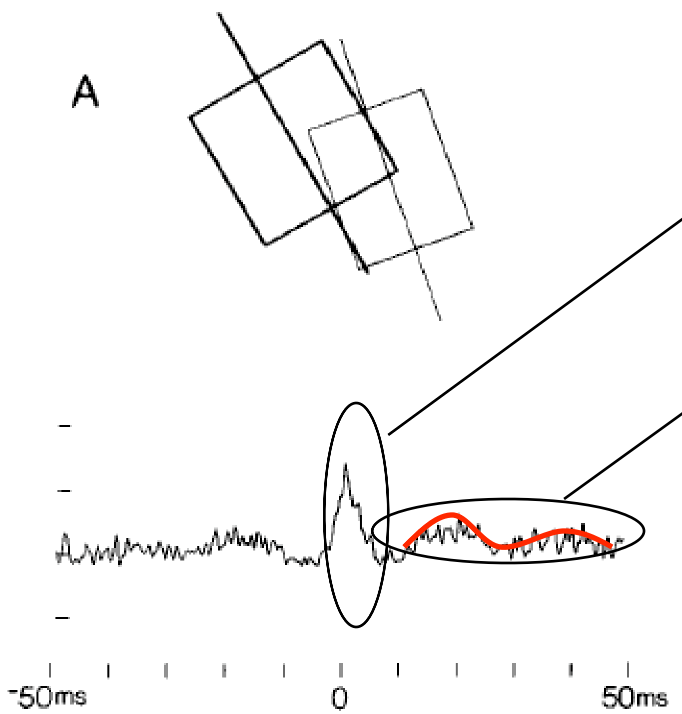
Overview of morning talks and tutorial

- This talk: *A* tutorial overview of signal processing methods for neural data
- Next talk: Data examples pertaining to tutorial talk
- Tutorials: *A*nalysis of individual data sets to illustrate the methods discussed in the two talks

Overview of this talk

- Quantifying auto and cross-correlations in time series using spectral measures
- Basic concepts: Sampling theorem, Nyquist frequency, DFT, FFT
- Time frequency resolution and the spectral concentration problem
- Multitaper spectral estimation
- Different methods for specifying point processes; point process spectra
- Singular value decompositions

Time domain correlation functions are popular in neuroscience to characterize correlations between neurons …



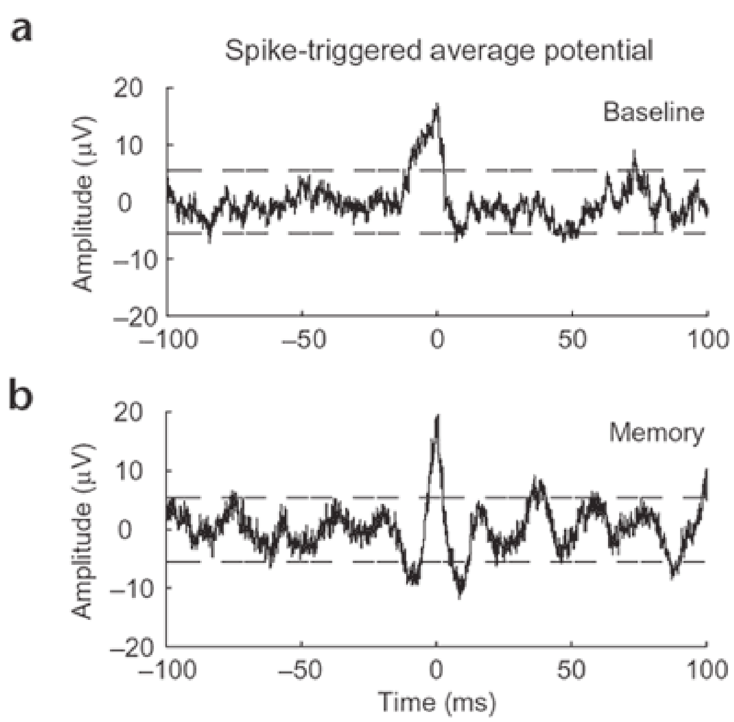
Daniel Y. Ts'o, Charles D. Gilbert, and Torsten N. Wiesel

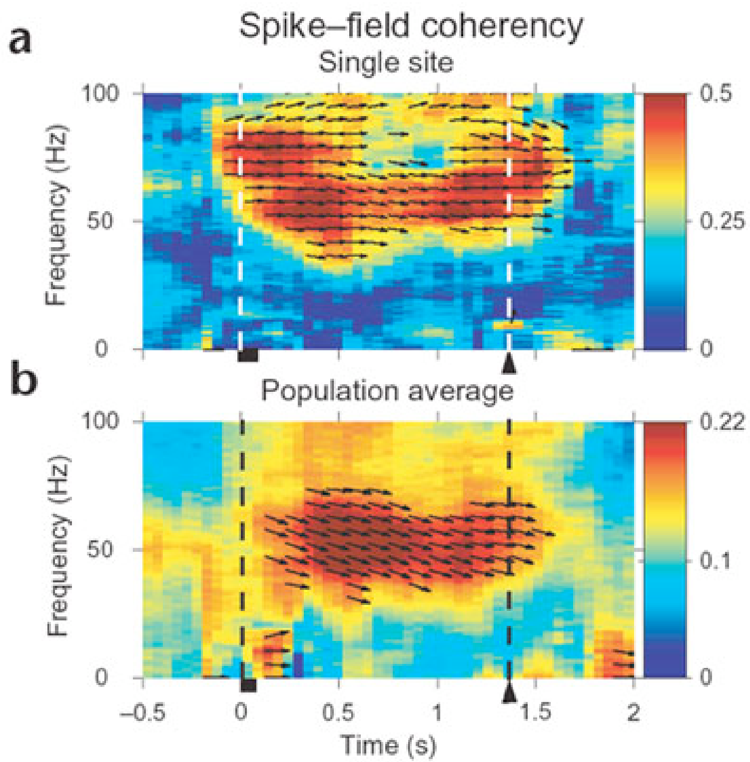
The Journal of Neuroscience
April 1986, 6(4): 1160–1170

- These can capture sharply peaked correlations ..
- .. But have difficulty detecting correlations distributed over time, such as those caused by quasi-periodic oscillations in the data
- Also note:
- (a) no confidence limits
- (b) difficulties in quantifying the "strength" of the correlations, or pooling across neurons, for anything but the central peak.

Real time spectrogram

- Characterization of temporal correlation patterns in neural signals using time dependent spectra
- Shows changes with state of arousal and finer cognitive modulations

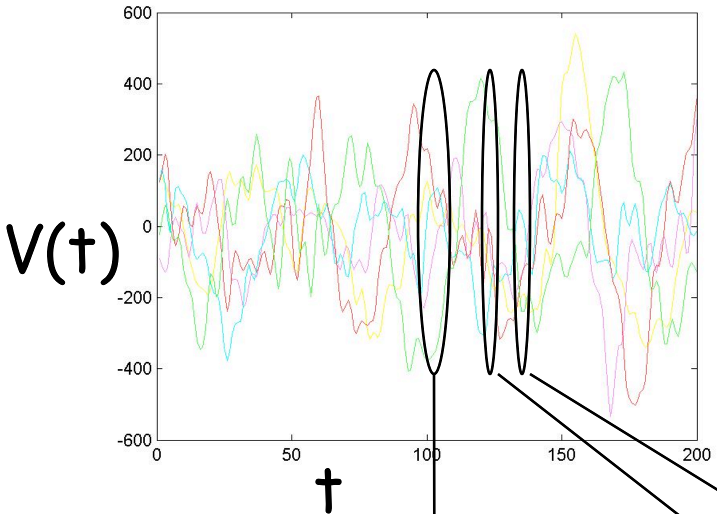


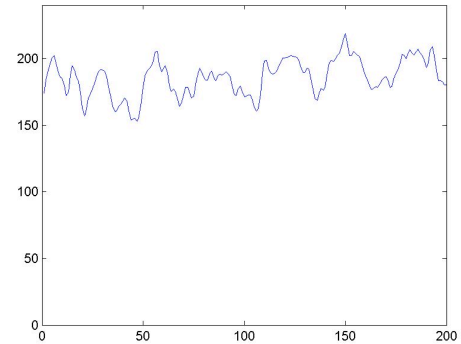


Time domain correlations show a peak at zero time lag, but an oscillatory part remains within the confidence intervals.

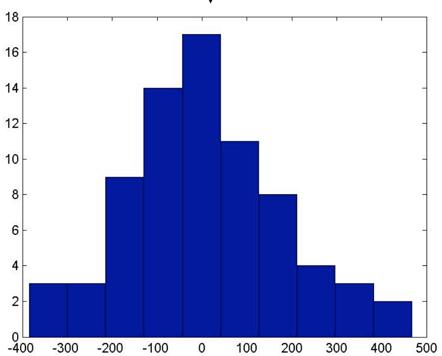
These oscillations appear to be significant on visual inspection, and in fact they do give rise to significant coherence in the time-frequency plane.

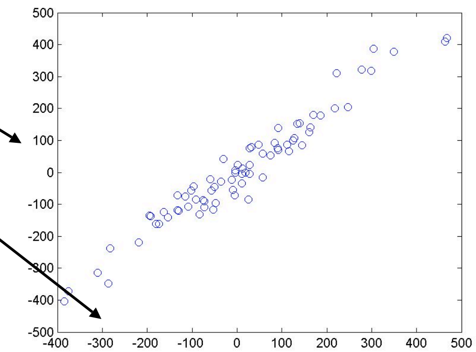
Let us look at some voltage waveforms (actual LFP recordings)…





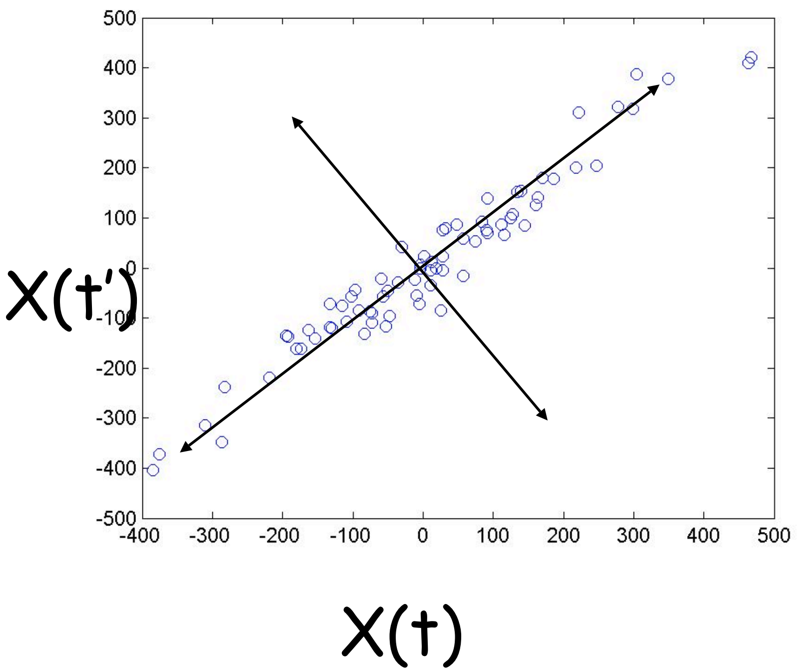
The variance is constant across time





At any given time, the voltages have a distribution, with a mean and a variance ..

… but nearby time points are correlated!



In the time basis, correlations between different points are nonzero ..

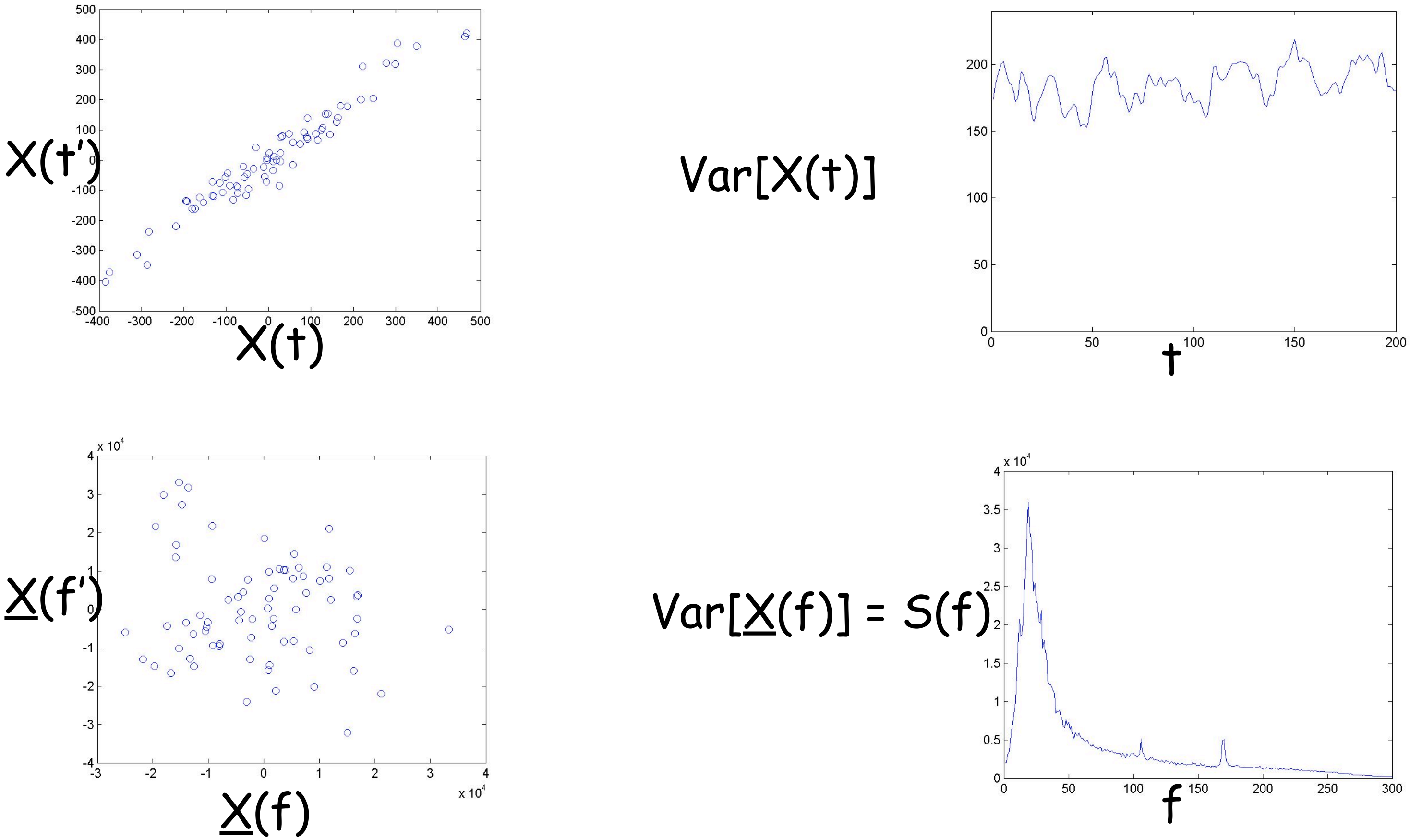
C(t)=cov[X(t+t'),X(t')] is nonzero

(assume the mean E[X(t)] has been subtracted)

Q. Can one go to a basis where the correlations are zero?

A. Yes! One can "rotate" the basis, and make the pair wise correlations vanish. For a stationary process, this is achieved by going to the frequency domain.

Cov[$\underline{X}$(f),$\underline{X}$(f')] = 0 unless f=f'



- For non-white processes, correlations between two time points C(t) is in general nonzero. Most processes of interest are not white.
- For stationary processes, correlations between two unequal frequencies is zero. Many processes of interest may be modeled as locally stationary.
- The variance at each frequency is given by the power spectrum S(f)

The continuous Fourier transform …

$$X(f) = \int x(t)e^{-2\pi ift}dt$$

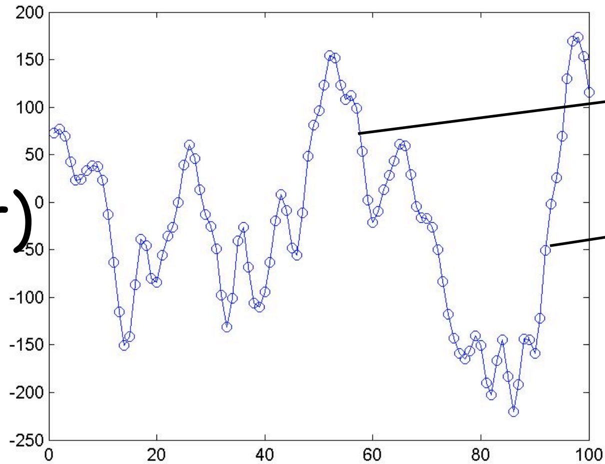
The autocorrelation function and the spectrum are related by Fourier transformation

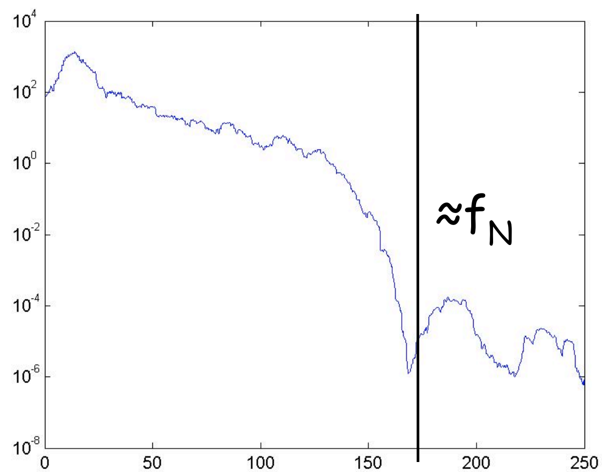
$$C(t) = < x(t+t')x(t') > - < x(t+t') >< x(t') >$$

$$< X(f)^{*} X(f') > = S(f)\delta(f-f')$$

$$S(f) = \int C(t)e^{-2\pi ift}dt$$

Sampling theorem, Nyquist Frequency, and all that …





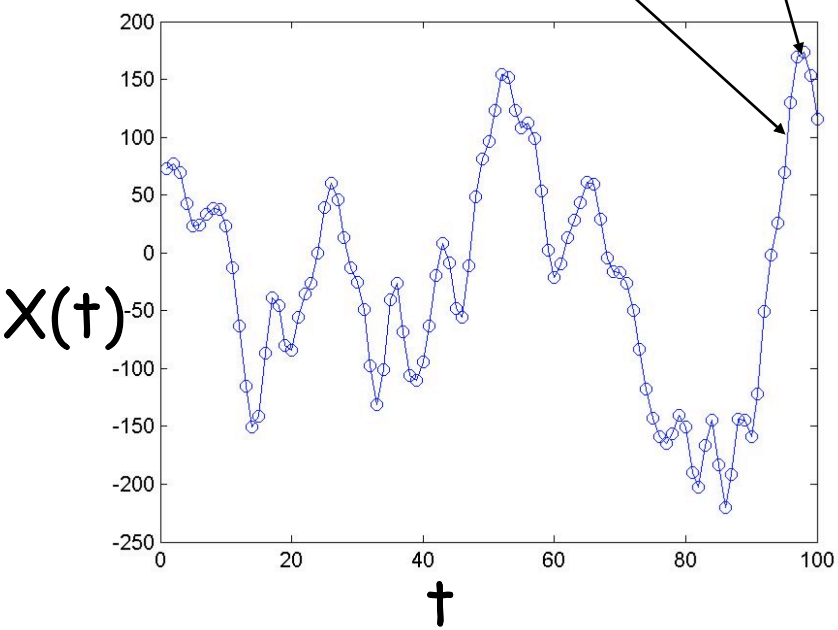
Approximately bandlimited at this frequency (note log scale)

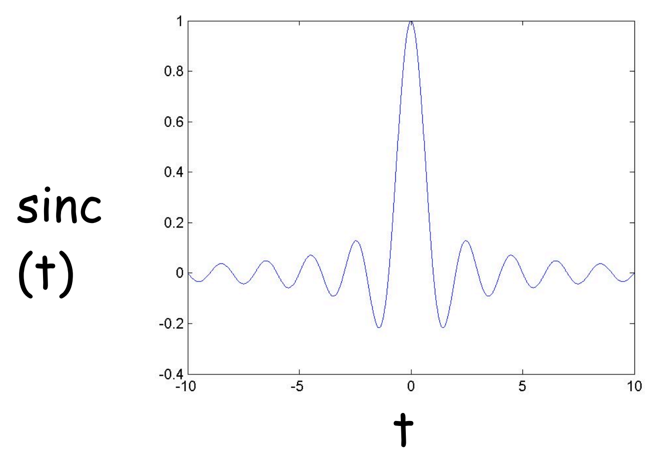
- Q. When can a continuous signal be fully characterized by discrete time samples?

- A. Under two conditions:
- (a) It is bandlimited (the spectrum is zero outside a finite interval $[-f_N, f_N]$)
- (b) The sampling frequency is larger than the size of this interval ($f_S > 2f_N$).

Sampling theorem tells us how to reconstruct the continuous time process from discrete time samples!

$$x(t) = \sum_n x_n \operatorname{sinc}[2\pi f_N (t - n\Delta t)]$$





Estimating spectra from data: tricky business

- The statistical theory of estimating spectra robustly from small samples poses nontrivial challenges.
- Blindly hitting the "FFT button" will not in general yield good results.
- Key problem: Time-Frequency uncertainty

FFT and DFT

The Discrete Fourier Transform is a continuous function of frequency obtained by transforming a discrete-time series

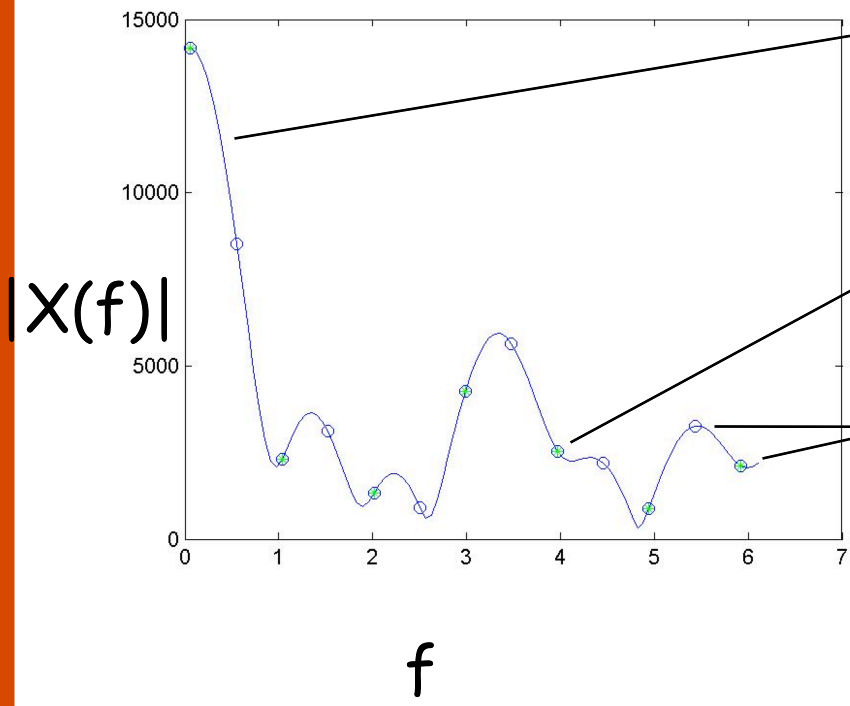
$$X(f) = \sum_{t=-T/2}^{T/2} x(t) e^{-2\pi i f t}$$

f here is continuous, and t is discrete (t=…,0,Δt,2Δt,..,nΔt,…)

The FFT is a fast algorithm to evaluate the DFT on a discrete frequency grid.

The resolution of the frequency grid is set by the amount of zero-padding used when evaluating the FFT.

DFT and FFT



- DFT (function of continuous frequency)
- FFT evaluated without padding
- FFT evaluated after padding by a factor of 2

Frequency resolution: The spectral concentration problem

If we had a discrete time series for infinite time, we would be able to evaluate its Fourier transform X(f) where

$$X(f) = \sum_{-\infty}^{\infty} x(t) e^{-2\pi i f t}$$

However, we only get finite segments of data (and if the process is nonstationary, then we may have to estimate spectra with even smaller segments). Therefore, we can only evaluate the truncated DFT $X_T(f)$

$$X_T(f) = \sum_{-T/2}^{T/2} x(t) e^{-2\pi i f t}$$

Dirichlet Kernel: Fourier transform or a rectangular window

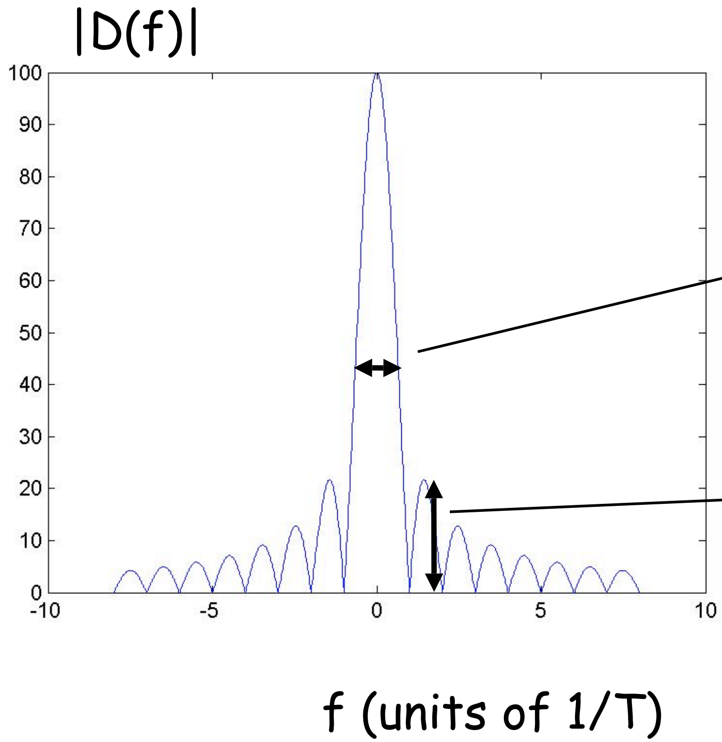
It can be shown that $X_T(f)$ is equal to $X(f)$ convolved with ("smeared by") the Dirichlet kernel

The Dirichlet kernel is the Fourier transform of a rectangular window.

$$X_T(f) = \int_{-1/(2\Delta t)}^{1/(2\Delta t)} D_T(f - f')X(f')df'$$

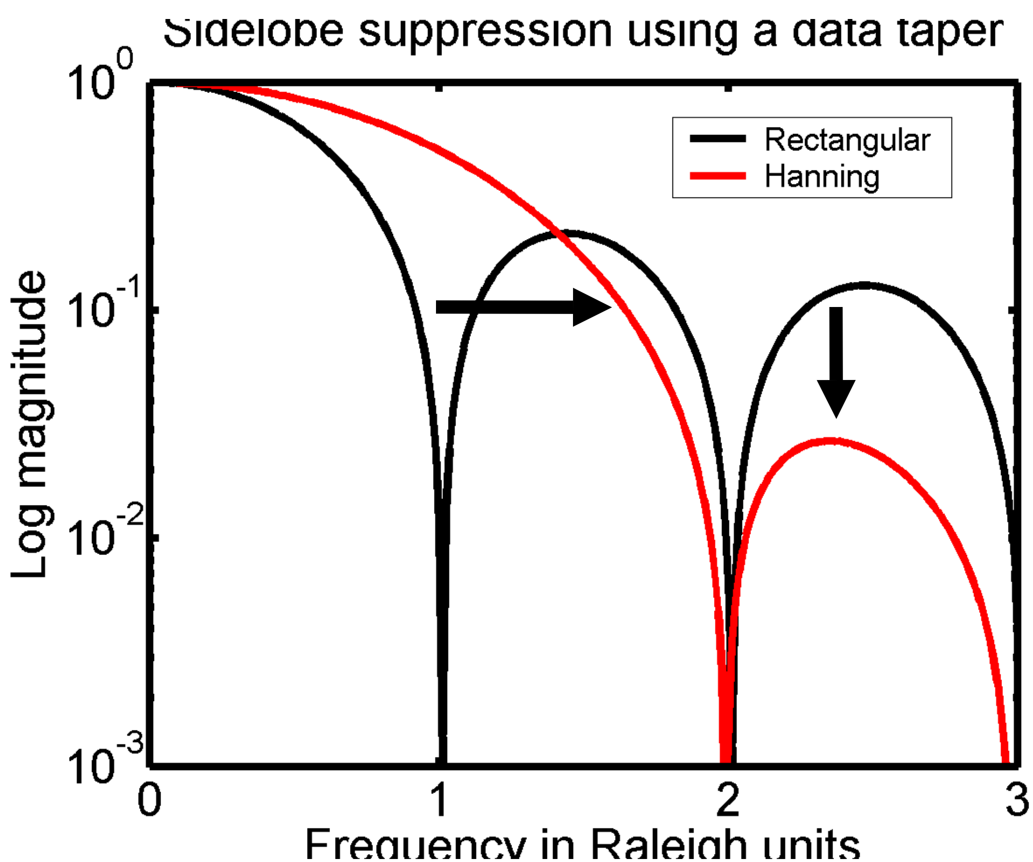
$$D_T(f) = \frac{\sin(\pi f T)}{\sin(\pi f)}$$

Narrowband and broadband bias



- Two problems caused by the finite window:
- (a) central lobe has finite width, $\Delta f=2/T$ ("narrowband bias")
- (b) Large side lobes: height of first side lobe is 20% of central lobe ("broadband bias").

“Tapering” the data with a smooth function (hanning, hamming, etc) reduces the sidelobe height, at the expense of the central lobe width … but:



- Are there “optimal” tapers?
- Tapering causes us to down-weight the edges of the data window (we lose data). Is there a way of recovering this information?
- These questions are elegantly answered within the framework

First, we consider the

Spectral Concentration Problem:

Find strictly time localised functions w_t , t=1..T,
Whose Fourier Transforms are maximally localised
on a frequency interval [-W,W].

This gives a basis set (Slepian functions) used for
Spectral estimation on finite time segments

Find functions w_t whose Fourier Transform U(f)

$$U(f) = \sum_{t=1}^{T} w_t e^{-2\pi i f t}$$

Are maximally concentrated in the frequency interval [-W,W]. To do this, maximise λ given by

$$\lambda = \frac{\int_{-W}^{W} |U(f)|^2 \, df}{\int_{-1/2}^{1/2} |U(f)|^2 \, df}$$

Maximising the spectral concentration ratio gives rise to an eigenvalue equation for a symmetric matrix

$$\sum_{t=1}^{T} \frac{\sin[2\pi W(t-t')]}{\pi(t-t')} w(t') = \lambda w(t)$$

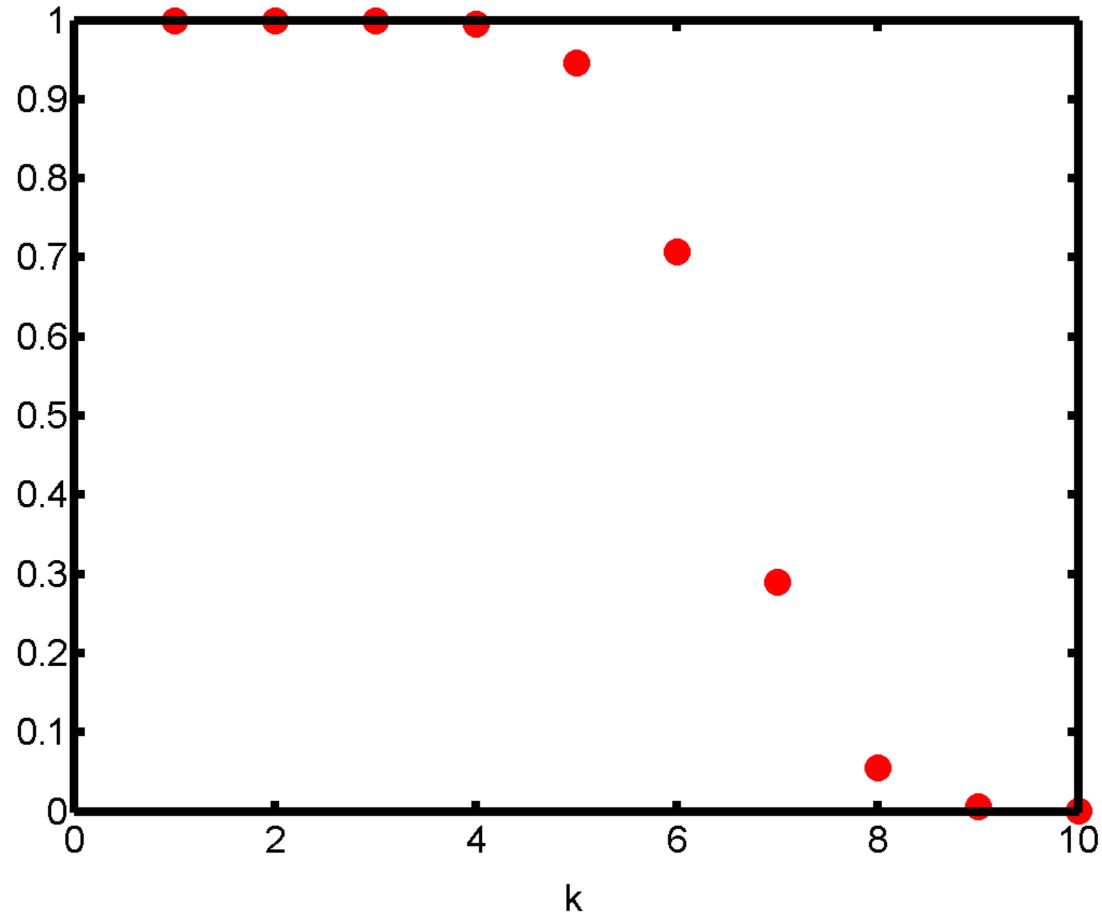
- Eigenvectors = Slepian functions

- Orthonormal on [-1/2,1/2] and orthogonal on [-W,W]

- K=2WT Eigenvalues are close to 1 (rest close to 0), Corresponding to 2WT functions concentrated within [-W,W]

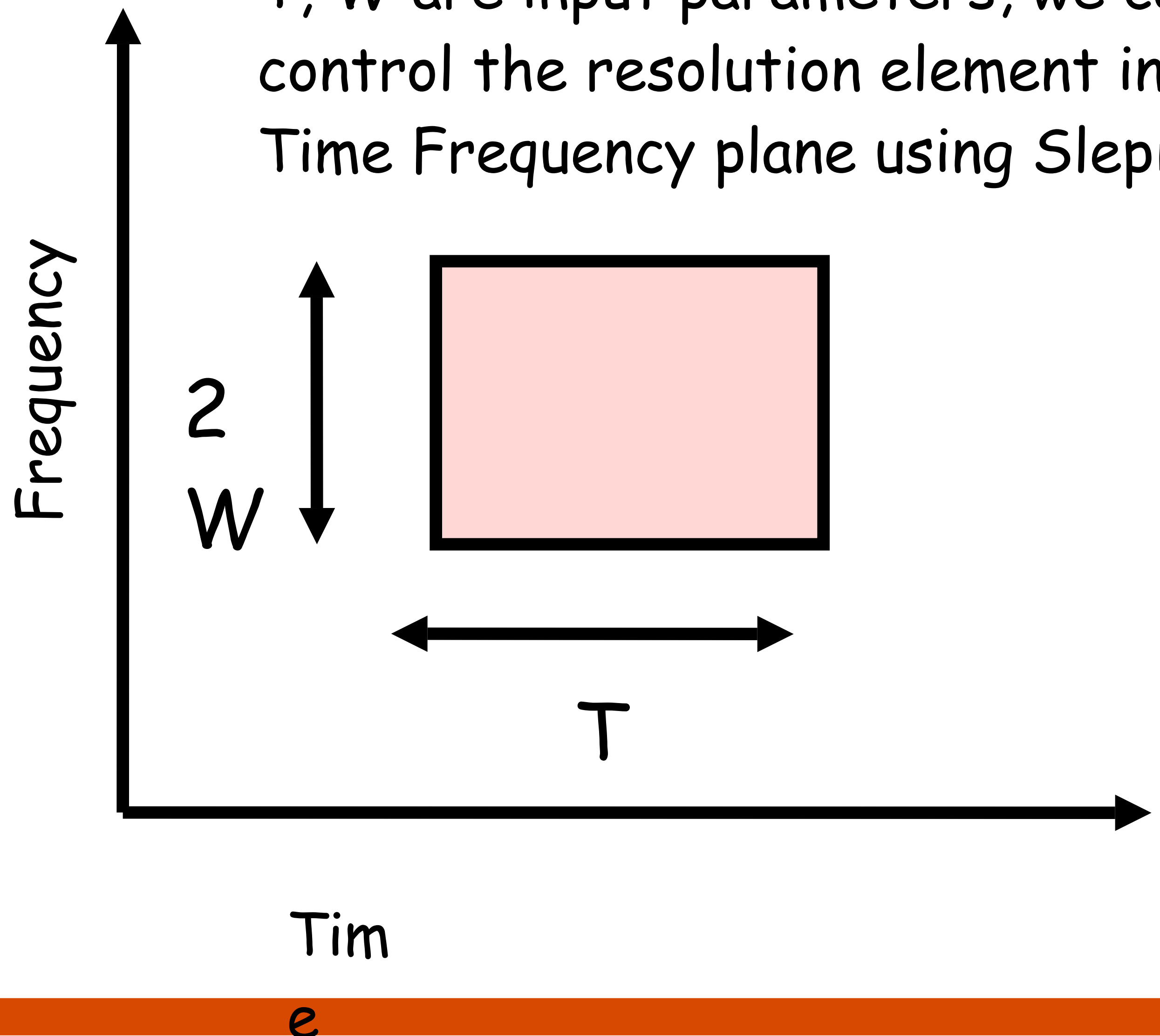
First 10 eigenvalues for 2WT=6

λ_k



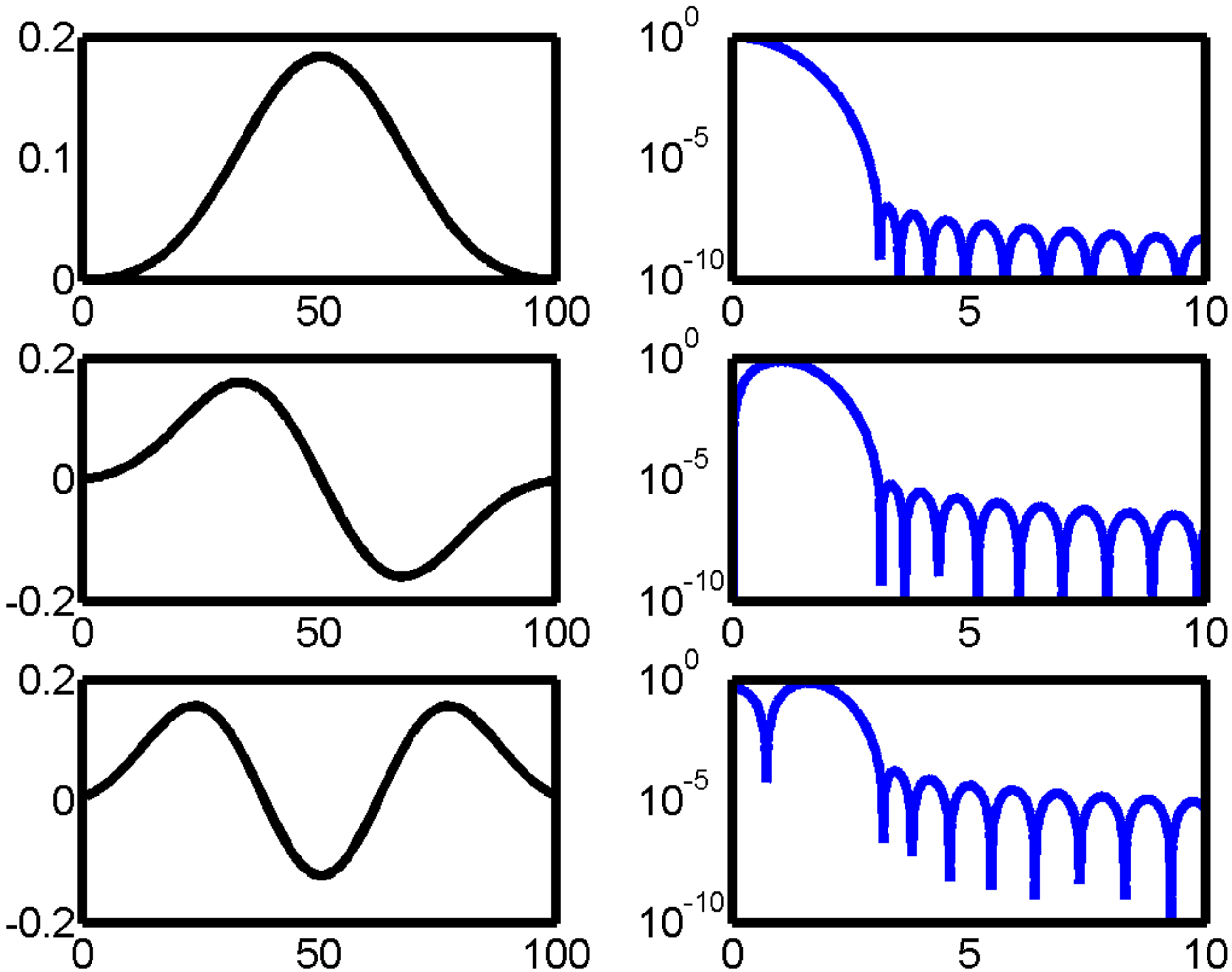
Λ_k is the power of the k^{th} Slepian function within the bandwidth [-W,W]

Approximately 2WT Slepian functions fit on this Time-Frequency tile. Since T, W are input parameters, we can easily control the resolution element in the Time Frequency plane using Slepians.

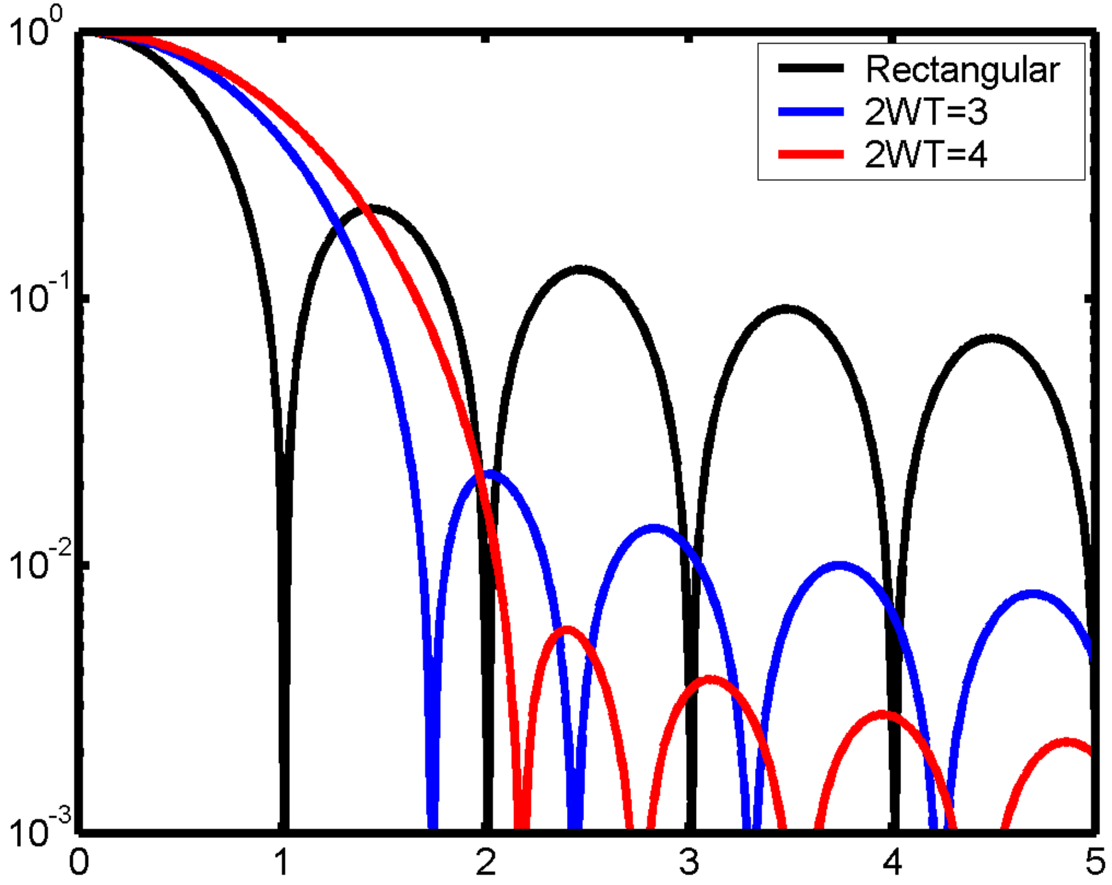


First 3 Slepians (2WT=6)

Corresponding Fourier power

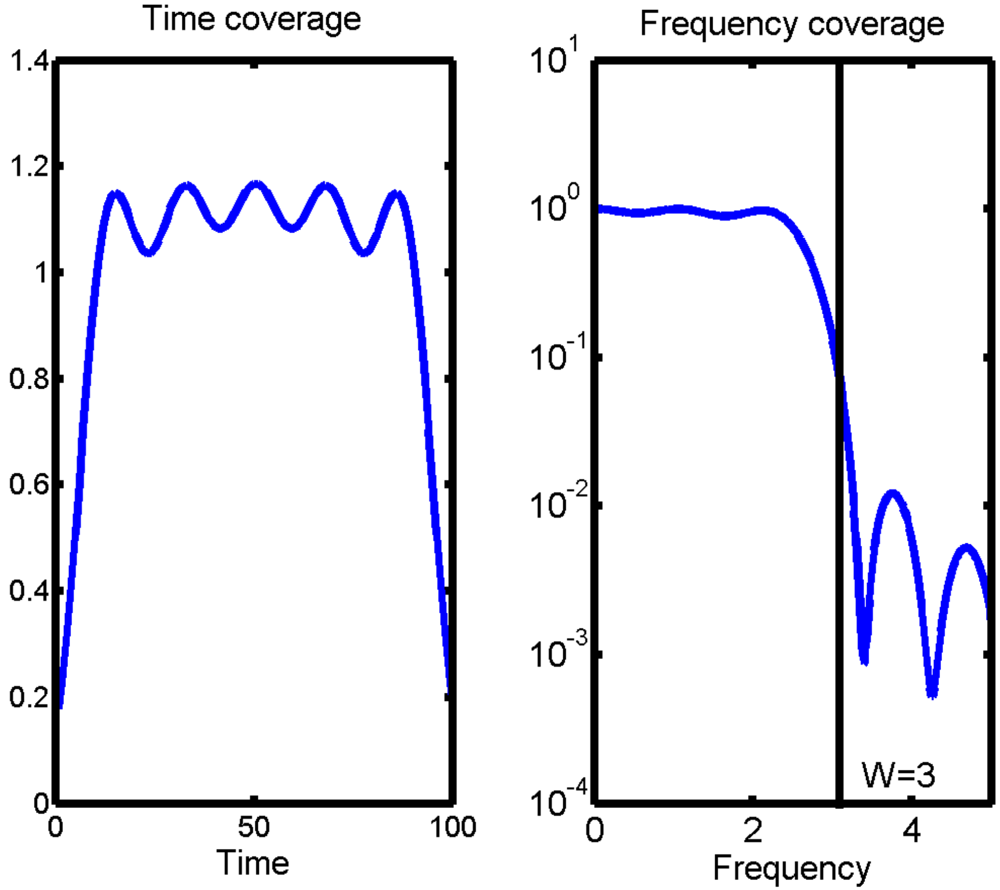


Slepian functions provide a systematic tradeoff
Between narrowband bias (central lobe width)
And broadband bias (sidelobe height)



Using multiple tapers recovers the edges of the time window (note the almost rectangular shape of the effective time-weighting function)

$$\sum_k w_k^2(t)$$



$$\sum_k |U_k(f)|^2$$

2WT=6

Multitaper Spectral Estimation:

$$S_{MT} = \frac{1}{K}\sum_{k=1}^{K} \|x_k(f)\|^2$$

where

$$x_k(f) = \sum_{t=1}^{N} w_t(k) x_t e^{-2\pi i f t}$$

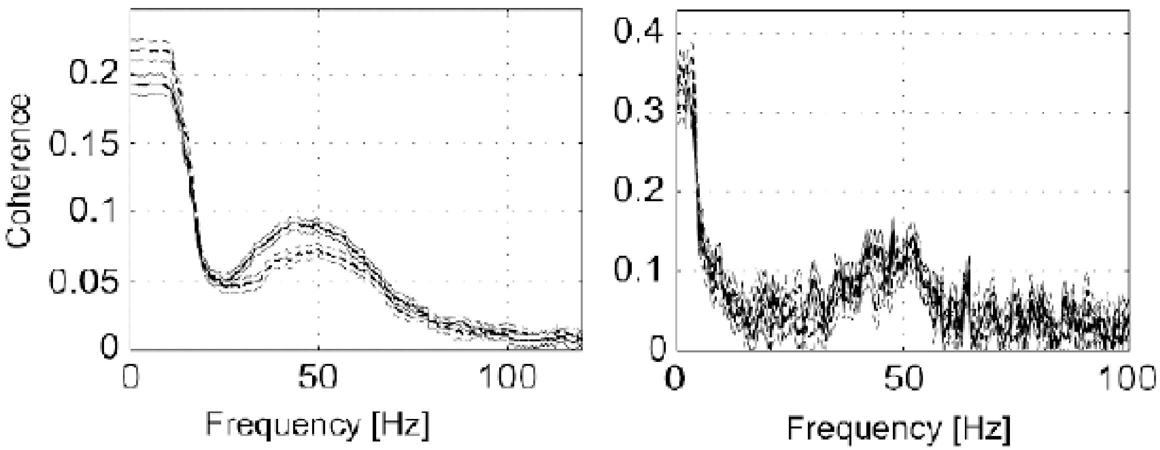
and $w_t(k) \quad ; \quad k = 1, \ldots, K$ are Slepian functions.

Cross Spectra and Cross Coherences between pairs of time series may be estimated in an analogous manner:

$$S_{XY}(f) = \frac{1}{K}\sum_{k=1}^{K} x_k^*(f) y_k(f)$$

$$C_{XY}(f) = \frac{S_{XY}(f)}{\sqrt{S_{XX}(f) S_{YY}(f)}}$$

An example: Comparing spike-LFP coherence between two conditions. Multitaper estimates with a large bandwidth compared with a narrowband estimate. Local confidence limits are indicated.



Smoothed estimates:
Confidence bands separated

Less smoothing:
Confidence bands overlap

Spectral Analysis for Point Processes

Different ways of specifying Point Processes

- The Configuration Probability (the joint distribution of all points)
- The Conditional Intensity (the probability of finding a point at a given time, conditioned on the past history)
- By specifying Moments of the process (or correlation functions)

(1)Specification in terms of Configuration Probability:

For all n, specify the joint probability of occurrence of all points, P(t1, t2, .., tn; n)=P(t1,t2,..,tn|n)p(n)

P()

Note: This configuration probability has to be given jointly for all n.

Example: Poisson Process

$$P(t_1, t_2, .., t_n \mid n) = \frac{1}{T^n}$$

$$p(n) = \frac{(\lambda T)^n}{n!} e^{-\lambda T}$$

Note: In this example, choosing p(n) to be other than the Poisson distribution would result in a non-Poisson process (An example of an application to spike trains: Richmond and Wiener, 2003)

(2) Specification in terms of Conditional Intensities:

Probability of occurrence of a point at a given time, given the past history of the process

$$\lambda(t \mid t_1, t_2, \ldots, t_n; n)$$

Given these ..

Λdt is the probability that a point occurs in the interval (t,t+dt)

For a Poisson process, the conditional intensity is a constant.

More generally, it is a random variable that depends on the sample of the point process.

Examples of applications to spike trains:
Brillinger et al, Brown et al.

(3) Specification in terms of moments of the process (analogous to specifying a univariate PDF in terms of the moments E(x), E(x^2), E(x^3), ..., E(x^n), ..)

The moments can be defined analogously to continuous processes:

$$\eta(t) = \sum_i \delta(t - t_i)$$

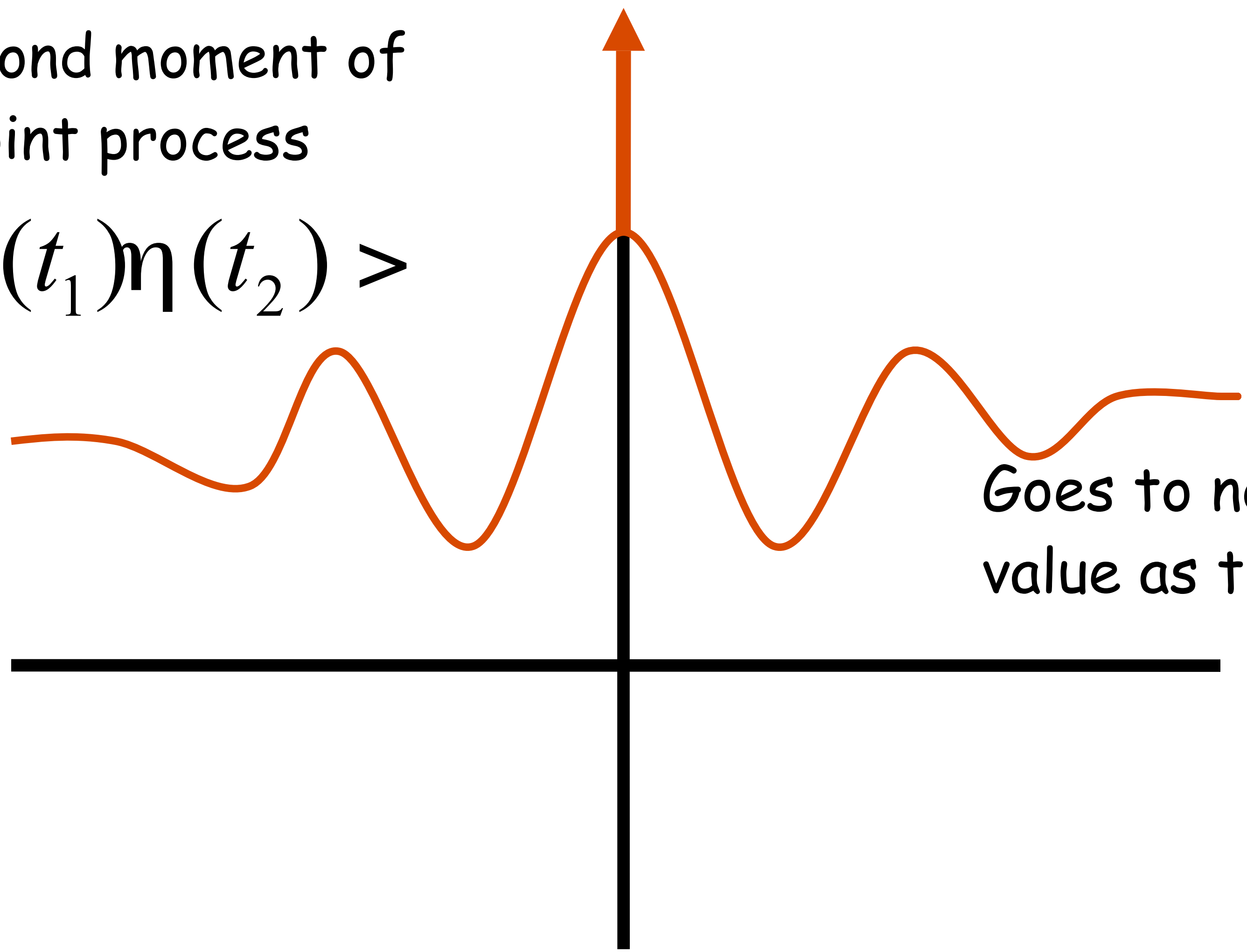
$$\text{First moment} = \langle \eta(t) \rangle$$

$$\text{Second moment} = \langle \eta(t_1)\eta(t_2) \rangle$$

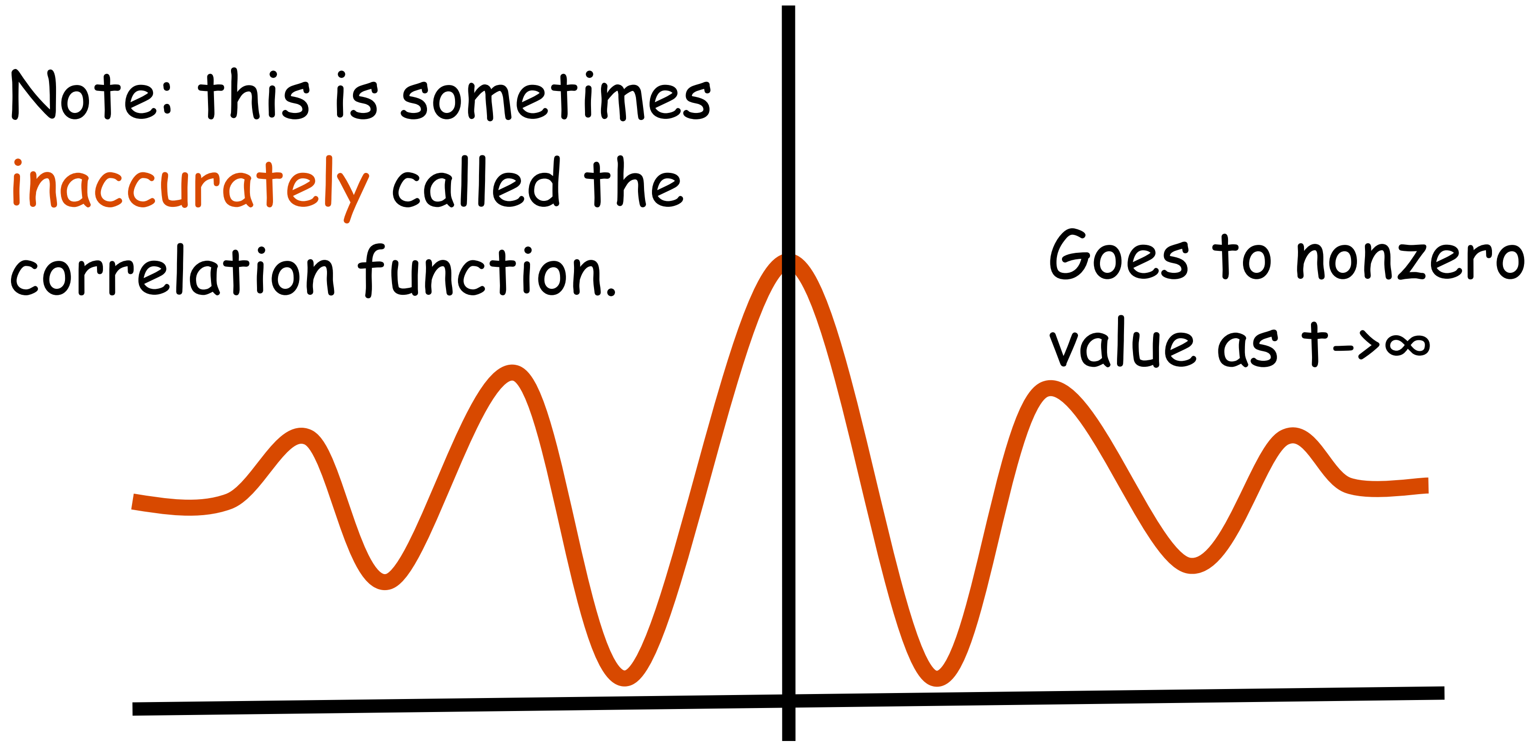
$$....\text{n}^{\text{th}} \text{ moment} = \langle \eta(t_1)\eta(t_2)..\eta(t_n) \rangle$$

Second moment of
a point process

$$<\eta(t_1)\eta(t_2)>$$



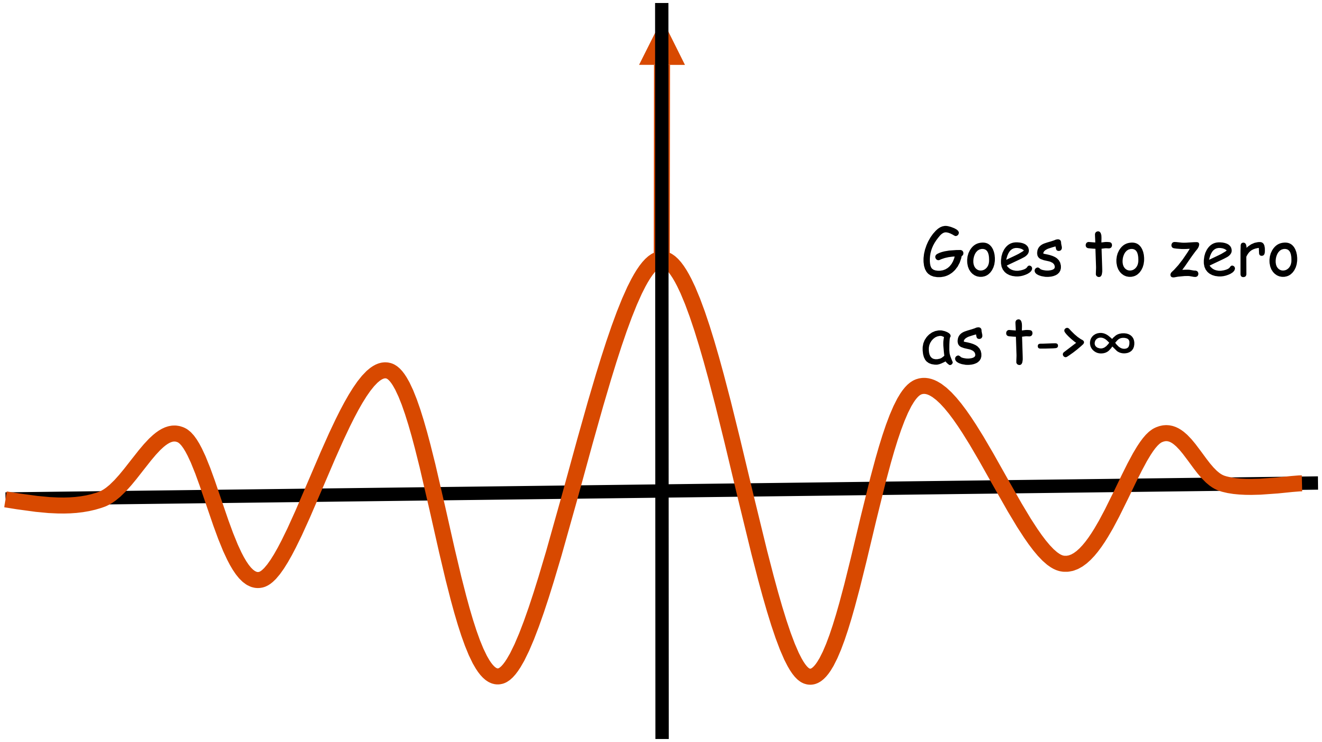
Product density: subtract the delta function



$$\rho(t_1,t_2) = <\eta(t_1)\eta(t_2)> - <\eta(t_1)>\delta(t_1-t_2)$$

$$Stationarity \Rightarrow \rho(t_1,t_2) = \rho(t_1-t_2)$$

Correlation function: subtract the asymptotic value



$$C(t_1,t_2) = <\eta(t_1)\eta(t_2)> - <\eta(t_1)><\eta(t_2)>$$

$$Stationarity \Rightarrow C(t_1,t_2) = C(t_1 - t_2)$$

The spectrum is given by the Fourier transform of the correlation function (as for the continuous process)

$$S(f) = \int dt C(t) e^{-2\pi i f t}$$

As f->∞, S(f)->R (the rate of the process)

$$\lim_{f \to \infty} S(f) = R = \lim_{T \to \infty} \frac{< N(T) >}{T}$$

As f->0, S(f) is given by the "Fano factor"

$$\frac{S(0)}{S(\infty)} = F = \lim_{T \to \infty} \frac{V(N(T))}{E(N(T))}$$

Example: For a Poisson process

$$\text{Second moment}: \langle \eta(t_1)\eta(t_2) \rangle = \lambda\delta(t_1 - t_2) + \lambda^2$$

$$\text{Two-point density}: \qquad \rho(t) = \lambda^2$$

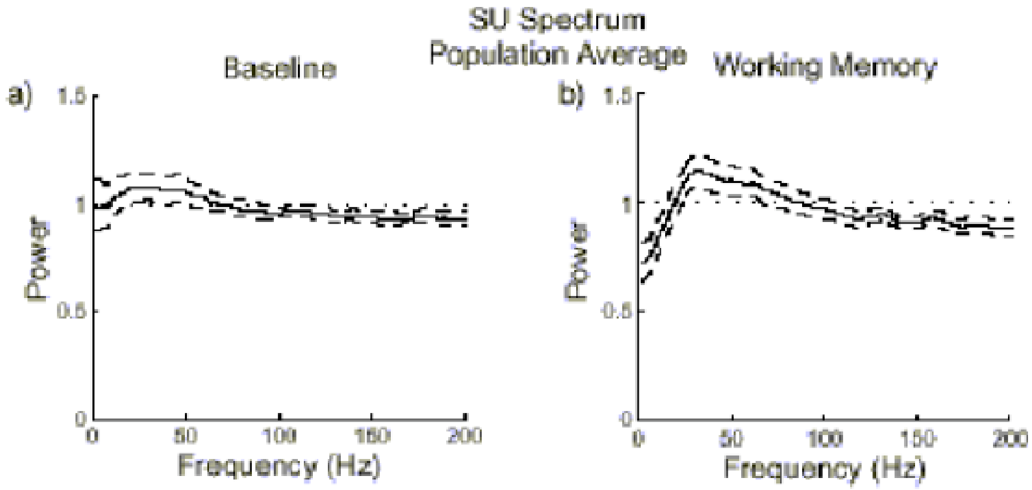
$$\text{Correlation function}: \qquad C(t) = \lambda\delta(t)$$

$$\text{Spectrum} \quad : \ S(f) = \int dt \lambda\delta(t) e^{2\pi i f t} = \lambda$$

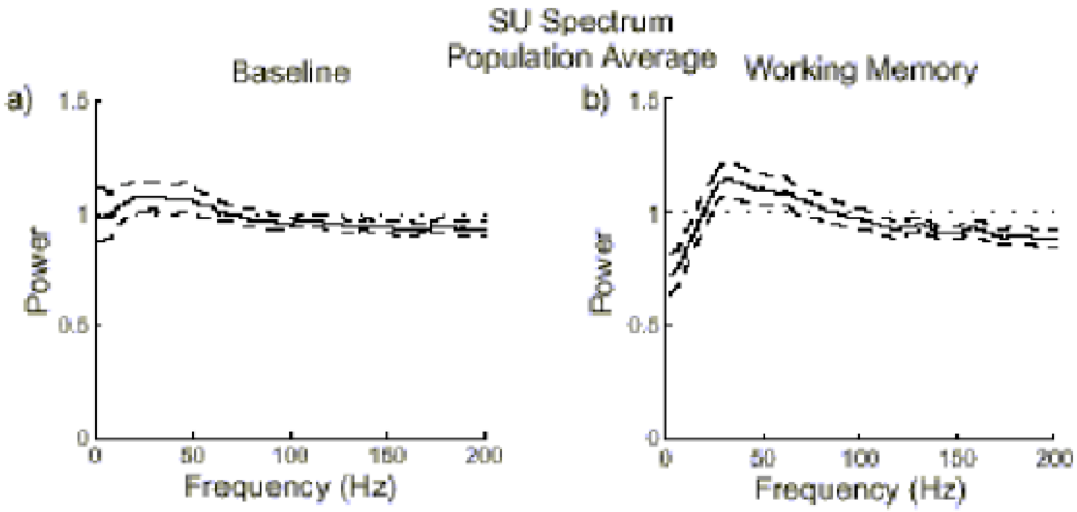
The spectrum is flat – the Poisson process is the point process analog of white noise.

Ruling out classes of processes:

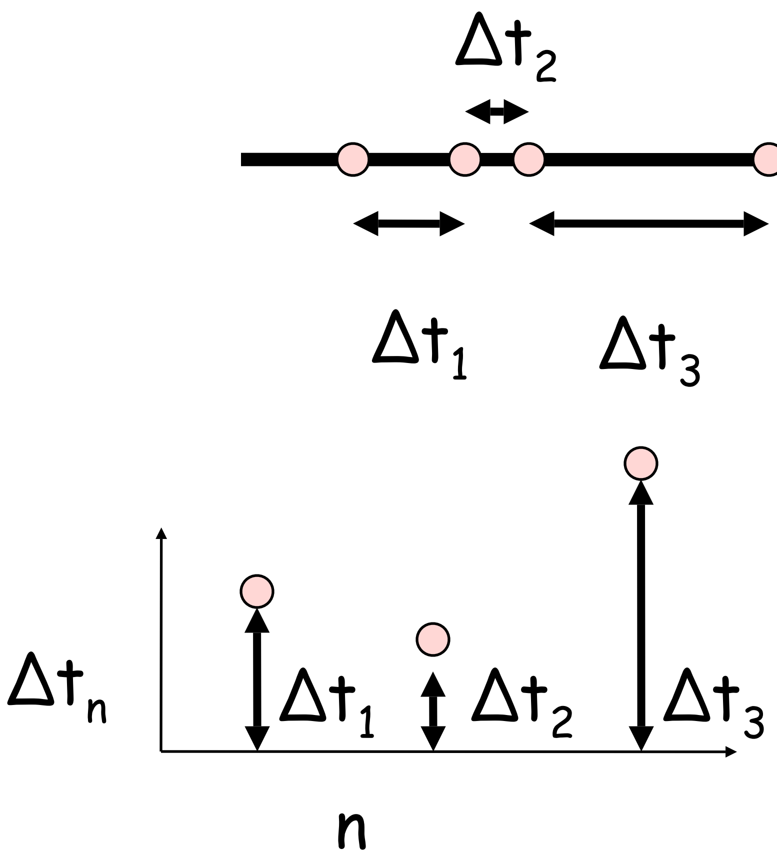
Rate Modulated Poisson: $S(f,t) = R(t)$. Therefore, if $S(f,t)$ shows frequency dependence then one can rule out a rate modulated Poisson.



Doubly Stochastic Poisson: $S(f) \geq \lambda = S(\infty)$. If the spectrum is anywhere less than its asymptotic value, doubly stochastic Poisson can be ruled out.



Interval spectrum: Spectrum of the process constructed by taking successive Inter Spike Intervals (ISIs)

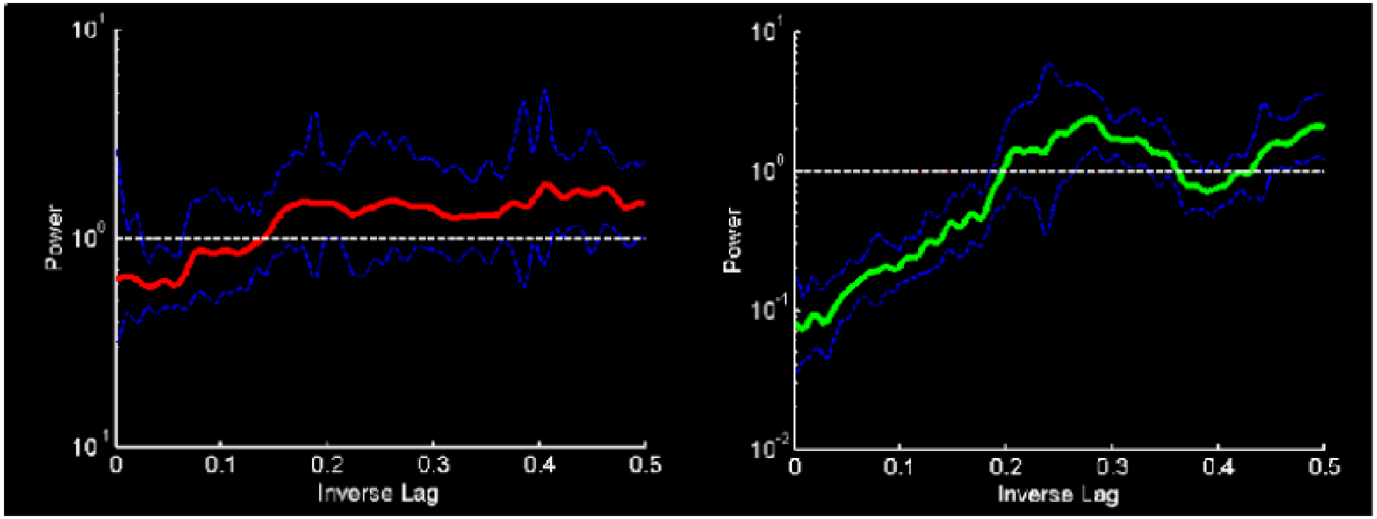


Estimate spectrum

$S_I(f)$

Renewal Process: since successive intervals are uncorrelated, the 'interval spectrum', ie the spectrum of the ISI process, is white. Renewal processes can be ruled out if the *interval spectrum* departs from a constant (or equivalently, successive ISIs are correlated).

Interval Spectra



Association between processes can be measured using the cross-coherence as for continuous processes

$$C_{xy}(f) = \frac{S_{xy}(f)}{\sqrt{S_x(f)S_y(f)}}$$

The zero frequency coherence gives the number covariation.

$$C_{ij}(0) = \frac{cov(N_i, N_j)}{\sqrt{V(N_i)V(N_j)}}$$

Some benefits of spectral measures:

•Unified framework for continuous and point processes

•Pooling across pairs easier (cf. Coherence magnitude)

•Smoothing in frequency reduces estimation variance

•Time-frequency methods account for nonstationarity

•Physiologically distinct sources often separate into distinct frequency bands, thus providing a natural method to separate 'signal' and 'noise'.

A variant on Spectral Analysis: Singular Value Decomposition.

Spectral Analysis $\leftrightarrow$ Factorisation of the temporal correlation matrix for single channels.

For multichannel data (EEG, MEG, fMRI), the SVD of the (channels x time) data matrix corresponds to a factorisation of the instantaneous correlation matrix.

All matrices (including rectangular ones) have a Singular Value Decomposition,

$$M = V \Lambda U^{\dagger}$$

Where V and U are unitary matrices and Λ is a diagonal matrix with non-negative entries called singular values. The columns of V and U are called left and right singular vectors, and satisfy the equation

$$\sum_j M_{ij} u_j = \lambda v_i$$

Counting: For a rectangular $p \times q$ matrix M, with $p \geq q$, there are at most q nonzero singular values.

$$M^{p\times q} = V^{p\times q}\Lambda^{q\times q}U^{\dagger q\times q}$$

For example, for image time series $I(x,t)$,

$$I(x,t) = \sum_n \lambda_n I_n(x) a_n(t)$$

Here n is given by the smaller of (number of channels or voxels) and (number of time points). The quantities $I_n(x)$ and $a_n(t)$ are called the spatial modes and temporal modes, or spatial singular vectors and temporal singular vectors.

The SVD is useful to analyse multichannel time series (including image time series) since in general, the singular value spectrum (a plot of the sorted singular values) shows a sharp rolloff, separating a 'signal subspace' from a 'noise subspace'. The 'noise tail' in the singular value spectrum can be analytically estimated for additive uncorrelated noise.

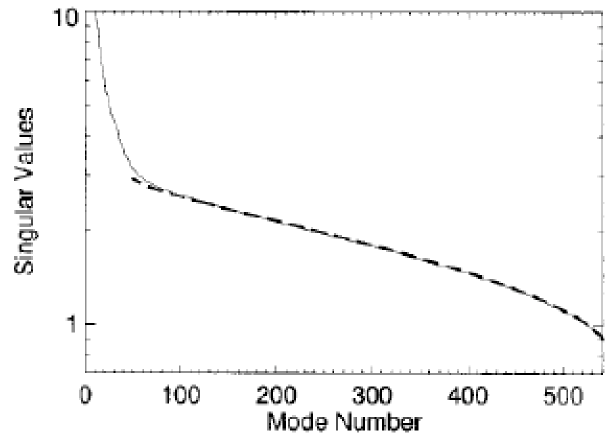


FIGURE 8 Sorted singular values corresponding to a space-time SVD of functional MRI image time series from data set ℰ. The tail in the spectrum is fit with the theoretical formula given in the text.

In general, individual modes have no physical or physiological meaning, except accidentally, and should not be selected by hand or ascribed such meaning. This is true for other decomposition algorithms including blind decomposition methods.

The exceptional case where a single mode can be interpreted is when there is only **one** large singular value above the noise background. *This often happens when dealing with a narrow frequency band.*

Time domain SVD modes do not in general segregate physiologically distinct sources. This can be rectified by confining the SVD to a narrow frequency band.

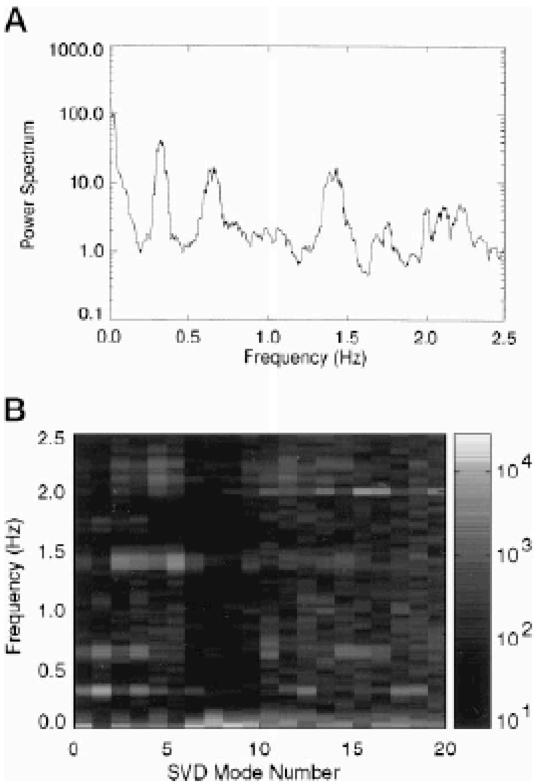


FIGURE 9 Spectra of first 20 principal component time series from data set €. (A) Weighted average spectrum; (B) image showing the spectra versus mode number. The log amplitudes of the spectra are color-coded.

SVDs are

- Useful in determining the correlation structure of the signals and the dimensionality of signal space.
- A data reduction or compression technique.
- An intermediate step useful for data conditioning and noise removal.
- A robust numerical algorithm available in MATLAB and other high performance numerical software.

The **Chronux** project

http://chronux.org

is an open source software project that aims to make **high quality signal processing tools for neurobiological data** widely available.

Current release includes a MATLAB toolbox containing a suite of time series analysis routines (includes time-frequency spectrum and coherence estimates for both continuous and point processes, local regression and local likelyhood estimates, confidence bands for estimates). There are also a number of GUIs included in the current release.

Collaborative development is encouraged, and feedback is welcome.