Java-DSP – Online Virtual Laboratories

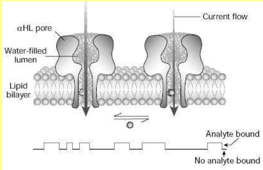
DSP Simulations

Extensions to other Disciplines

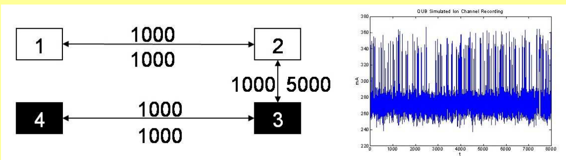
Sponsored by NSF Awards 0817596, NSF-DUE-CCLI-080975
NSF Program CCLI Phase 3 Award Started Apr. 2008 – Apr. 2013 involves 8 universities
Also core software used in an NSF CRCD 2004-2006

Classifying Signals from Ion-Channels

Basic Structure of an Ion-Channel



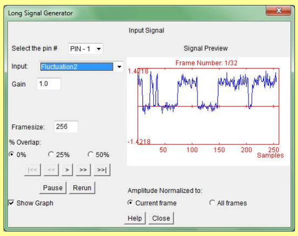
Ion Channel Signals



- We have used the QUB scientific package to generate synthetic data.
- Left panel shows a sample 4-state Markov model used for generating data and a sample trace is shown in Right panel.
- We constructed models to simulate responses of two highly similar analytes which closely resemble the authentic data. Utilizing multiple recordings, an input data matrix of dimension 400×10000 is formed by extracting four 10,000 point sequences of one second duration from each recording for a total of 50 input files for each analyte.

Java-DSP and Ion Channel

The Long signal generator block in J-DSP has time series data obtained from fluctuations1 in the conductance of two different channels.



Parameters of an Ion Channel

Single channel kinetics

Suppose that a channel can exist in only two states, open and closed;

$$\text{closed} \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} \text{open}$$

where, the rate constant for open to closed transitions will be called β, and α will be the opening rate constant $p = \alpha/(\alpha+\beta)$

Then, the simple spectrum and autocorrelation function are given by the following equations:

$$S(f) = \frac{2\gamma p(1-p)/(\alpha+\beta)}{1+(2f/(\alpha+\beta))2}$$

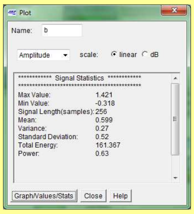
$$C(T) = \gamma p(1-p)e^{-(\alpha+\beta)t} \qquad p = \alpha/(\alpha+\beta)$$

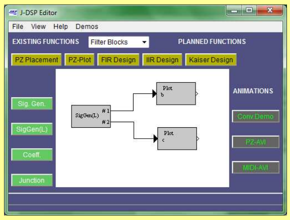
where, p is the steady state probability that the channel is open.

Noise Characterization

Statistics of Fluctuations

The statistical characteristics of These fluctuations of two channels are not identical, but rather reflect the physical process.

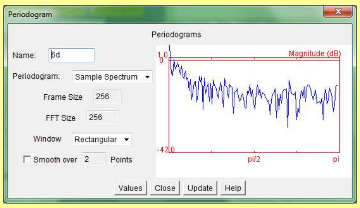


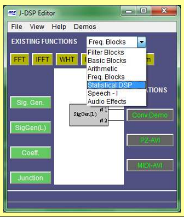


Spectral Analysis

Spectral Analysis

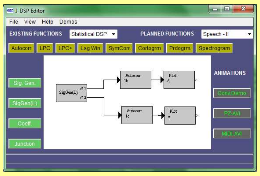
The statistical characteristics of These fluctuations of two channels are not identical, but rather reflect the physical process. Just as exponential functions arise in physical relaxation process, so does a particular spectral form called the simple spectrum.



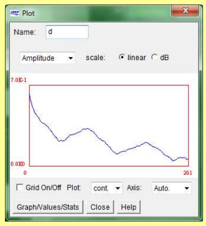


Autocorrelations

The spectrum is one of the two commonly used ways for characterizing fluctuations; the other is autocorrelation function. The autocorrelation function specifies over what time periods the noise is correlated. If the fluctuations are slowly varying at a particular time the chance of predicting right at a later time is high and vice versa.



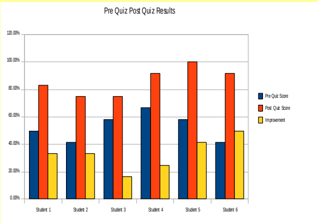
Block Connections for finding plotting Autocorrelations.



Exercise Evaluation

J-DSP on-line labs helps better understand and visualize the Fluctuation data of Membranes and analyze the various statistics.

Understanding of the concepts of spectral density and dependence on variance of fluctuations in membrane conductance is enhanced by the J-DSP Exercises.



The general concepts of using Autocorrelations in signal analysis is clear by performing a J-DSP simulation

Machine learning Approach

Walsh-Hadamard Transform (WHT)

The WHT is able to represent signals with sharp discontinuities more accurately using fewer coefficients.

For a given window size *N, it was determined that 20% of the WHT coefficients represents 90% of the signal energy.*

Thus by discarding the coefficients that do not contribute significantly to the signal energy, the size of the dataset was reduced by 80%.

Even after WHT is performed, further dimensionality reduction is required on the dataset. For example, *N=4096, the size of the transformed dataset is 400×819. It is likely that many of the selected coefficients are highly correlated and there is scope for further compaction.*

Feature Extraction

Principal components analysis (PCA)
It performs a linear mapping of multidimensional data to a lower dimensional space while retaining as much as possible of the data variability.
It was determined that that the first 10 components account for 99% of the total variance. Thus we project the data on the bases represented by the 10 principal components.
Now the dimension of the dataset is 400×10.

Training and Testing data
The dataset thus consists of 400 vectors. The transformed dataset is randomly permuted and partitioned into a training set of 200 vectors and test set of 200 vectors.
To compensate for the small size of the dataset, *m-fold cross-validation was used for model selection*

Classification Results

The following algorithms were used for classification:

Multi-Layer Perceptron (MLP)
Linear/Quadratic Discriminant Analysis (LDA/QDA)
Learning Vector Quantization (LVQ)
Radial Basis Function Networks (RBF)
Support Vector Machines (SVM).

Algorithm Used	Classification Performance (%)		
	N=1024	*N=2048*	*N=4096*
MLP	*71.0*	*73.0*	*81.5*
LDA/QDA	*64.1*	*72*	*78.0*
RBF	*67.0*	*71.5*	*76.5*
SVM-RBF	*69.0*	*74.5*	*80.5*
SVM-Polynomial	*66.5*	*72.5*	*80.0*

MLPs gave the best performance for all frame lengths.

Thank you …..