Breaking the Brain's Secret Code

Nathan S. Hicks

Prof. Larry Sorensen, Advisor

University of Washington
Department of Physics – MS Degree Project
June 2019

TABLE OF CONTENTS

TABLE OF FIGURES

ABSTRACT

This research paper targets the identification of Utah Electrode Array dataset patterns by using various analysis techniques from classical statistics to other non-traditional derivations. From a classical standpoint, electrode voltages of the Utah Array are verified as normally distributed, and multiple heatmap plots are developed to visualize the related electrode voltage and signal cross-correlation results. Through sequential matrices of plots and via animated heatmap plots, both false color and grey color scales are leveraged for different purposes. For the false color plots the primary audience is that of human interpretation, whilst the latter is intended for computer vision analysis (future work). In this regard, various Look Up Table transformation functions are at first evaluated and chosen to both improve the contrast of these heatmap visualizations, and to prevent oversaturation of the plotted results. Beyond this, scatter plots are developed as a fundamental building block for demonstrating pair-wise electrode dependencies, and are the primary constituents of a covariance matrix. For such a matrix, all 81 pair-wise electrode dependencies of the Utah Array are visually plotted, that results in multiple rectangular patterns of varying shade. Another approach used, comparing cross-correlation lag peak magnitudes at idealized lags achieves similar results. Here, the electrode cross-correlation differences are classified according to the more correlated of the electrodes, where several electrode voltage heatmap plots are given as direct evidence of such a result. In the end, recommendations are given for future work, suggesting that machine learning and computer vision analysis of the existing grey scale heatmap plots be undertaken as a priority.

INTRODUCTION

Achievements within the field of brain research are advancing rapidly every day toward deeper self-awareness. Gaining such introspective knowledge requires specialized medical monitor devices as developed across history, consequently entrusting scientists as knowledge-bearers of highly detailed studies upon human neural network activities. In modern times, the electroencephalograph (EEG) device, the electrocortocogram (ECoG) device, the magnetic resonance imaging (MRI) device, and the Functional MRI (fMRI) device are most commonly employed. In extreme circumstances, a Utah Electrode Array (UEA) is necessary, and in comparison with an EEG where electrodes are positioned external to the skull, the UEA requires

implant directly to the human cortex. This type of implantation positions the electrodes directly amidst the cortical neural networks and therefore provides improved resolution of intra-cortical synaptic transmissions of individual neurons [1]. These resultant synaptic voltage recordings, saved as data, provide a lower signal-to-noise ratio (versus an EEG) and the brain research field is subsequently enhanced in more meaningful ways. Indeed, substantial volumes of UEA research exist today, particularly with respect to analysis using traditional statistical methodologies. In this paper, an UEA dataset analysis is treated with traditional statistics prior to employing other novel (non-traditional) analytic methodologies.

Henceforth, this research paper intends to discover neuronal synaptic transmissions that reflect brain activity across the UEA for observable pattern detection that may not otherwise be self-evident. Several prior studies in this regard exist from within the University of Washington Department of Physics graduate studies program. Primarily, these prior works have sought conclusions using statistics and correlations of signals for UEA pattern identification, for human observable conclusions. This paper similarly re-affirms such prior statistical works using the same dataset as did all prior research studies therein, for evaluating scatterplots, probability densities, electrode correlations, and the like. Thereafter, this paper builds upon new novel implementations: animated plots such as histograms and heatmaps for example, and consideration is given to cross-correlation lag peak difference plots, and point-slope difference plots alike. The dataset considered here is that of a medical patient whom is at rest, and no pre-defined patient activities of specific interest exist. Most certainly, this lack of defined features (e.g. an absence of commanded body movements) makes for a challenging research problem, particularly if the desire is to literally, 'see someone's thoughts'. As such, the dataset of interest to this paper was derived from an UEA located at the right temporal lobe of a male epileptic patient. The patient was ensured awake and restful throughout the data recording session, lasting between the hours of 10 AM and 6 PM. The raw UEA voltages were amplified and then filtered with both a high-pass and a low-pass filter of a 0.3 Hz Butterworth and a 7.5 kHz 3-pole Butterworth, respectively, at a sample rate of 30 kHz. Finally, during data acquisition, these same electrode voltages were referenced to a subdural / epidural surface at the cortical implant site

[1]. The dataset used for analysis contains only 81 of the 100 UEA electrodes due to equipment faults in 1 of the rows and 1 of the columns.

BACKGROUND

THE NEURAL NETWORK

Dating back to the 18th century [2], scientists have aimed to discover the complexities that constitute homo sapiens through studies of the human brain. As a complex network of neuronal activity, synaptic transmissions within the cortex (Figure 1) are today understood to function via electrochemical principles upon which sodium and potassium ions exchange across a neuron's cell membrane. Protein structures acting as ion channels then respond linearly to these ionically induced voltage potentials until a threshold is reached [3], whereupon the preceding synaptic transmission is halted. This dynamic process outlines the neuron-to-neuron communications sequence. This neural network of activities constitutes the outer layer of the brain itself, beneath the boney cranium, called gray matter. With a depth of only 2-4 mm, the term gray matter is used to differentiate itself with white matter, that is positioned just beneath the prior [4], and consists primarily of neural axons acting as communication pathways that reach into the inner portions of the brain (Figure 2) [5].

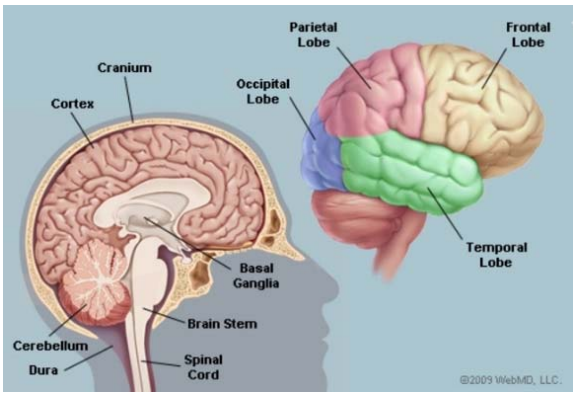

Figure 1 – Primary areas and structures of the human brain.

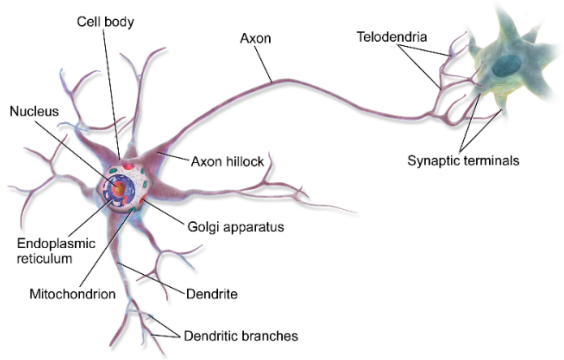

Figure 2 – The standard structure of a neuron of the human brain.

The combined classifications of both gray and white matter is referred to as the cerebrum, where the overall partitioning of the brain is routinely divided into 4 sections: temporal lobe, occipital lobe, parietal lobe, and the frontal lobe (see Figure 1). Lengthy consultations and

problem diagnosis are a pre-requisite for consideration of where to locate a Utah Electrode Array implantation within these divisions.

THE UTAH ELECTRODE ARRAY

As previously stated above, the Utah Electrode Array (UEA) is a cortically implanted device. The standardized UEA consists of a 10x10 square matrix of spatially located electrodes of 1.5-mm lengths and 4 µm diameters each. The UEA is constructed with a 0.2-mm thick silicon base, platinum electrode contacts, and glass dielectric for electrode isolation. The footprint of the unit is a mere 16-mm^2, where the inter-electrode spacing approximates the columnar inter-cortical spacing of 0.4-mm [6] [7]. The electrode tips themselves are sputter coated with titanium / tungsten and platinum, that facilitates the ionic-electric transduction process of data collection. Therefore, only the tip of the electrode is actively measuring the underlying voltage fluctuations. A specialized device constructed in this way, using a square geometry for electrode placement, permits specific spatial multi-electrode analysis techniques from the resulting voltage recordings. Figure 3 below shows the delicate construction of the UEA in close-up detail. Such electrodes do fail over time, and this could explain the initial dataset selection of only 81 electrodes from the raw dataset, as mentioned above, as these showed reliable data. Such outcomes might occur if the initial UEA placement is not fully interfaced to the brain, perhaps if the cortex is not locally flat [6]. Equally, after a 2-3 week implant period, when the immune system transitions to the chronic phase, the device begins to encapsulate with fibrous growth, and this will cause challenges to reliable data collection as well [8].

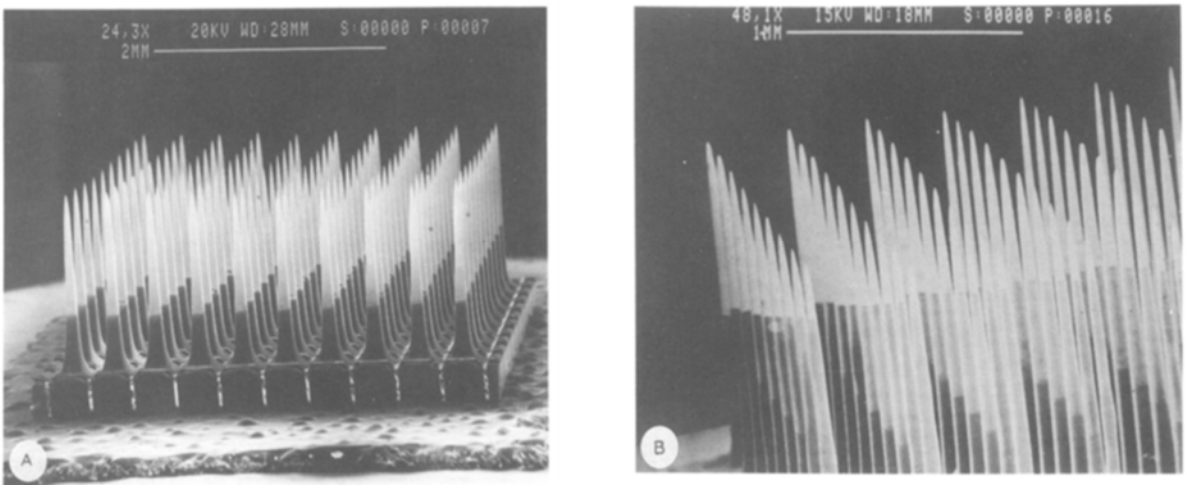

Figure 3 – Scanning electron micrograph of the cortically implanted Utah Array electrodes. The images here are platinum coated electrode tips.

ELECTRODE SIGNALS

With typical inter-neuron spacing between 500-µm and 2-mm [7], each electrode of an Utah Array receives its neural information from a small localized population of neurons located nearby the titanium / tungsten coated surface of its tip. This information stream of localized brain activity is captured and stored via an electrodes' voltage variances across time, and measured in units of millivolts, mV. For this research paper, data was originally sampled at 30 kHz and then reduced to a sample rate of 1 kHz, using a unit of measure in milliseconds, mSec [4]. Therefore, as a first inspection upon the Utah Electrode Array dataset of 81-electrodes, Figure 4 below plots the time series of neural activity for a single electrode for 1-million data points, that is, a 1,000-second time window.

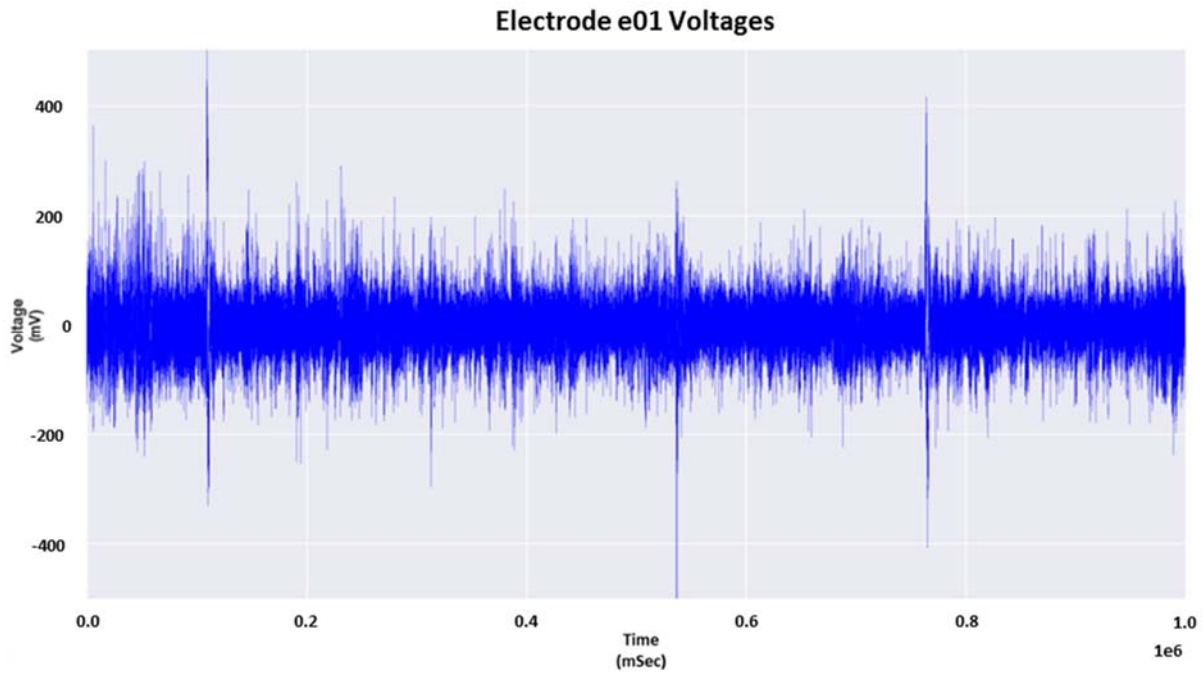

Figure 4 – Time series of raw dataset voltages for electrode 1, depicting its signal characteristics across 1,000-seconds.

Aside of the many random peaks among the electrode inputs, visual inspection of this plot quickly shows that the electrode maintains a mean of 0 mV throughout its period of activity. Therefore, the data recorded from electrode 1 of Figure 4 is that of a Gaussian distribution, as will be shown again in a later section of this paper. Gaussian distributions permit a traditional statistics approach to data analytics in that normal distributions and scatter plots, for example, can be assessed. In this same way, Figure 5 expands upon the entire Utah Array time series plots of all 81 electrodes.

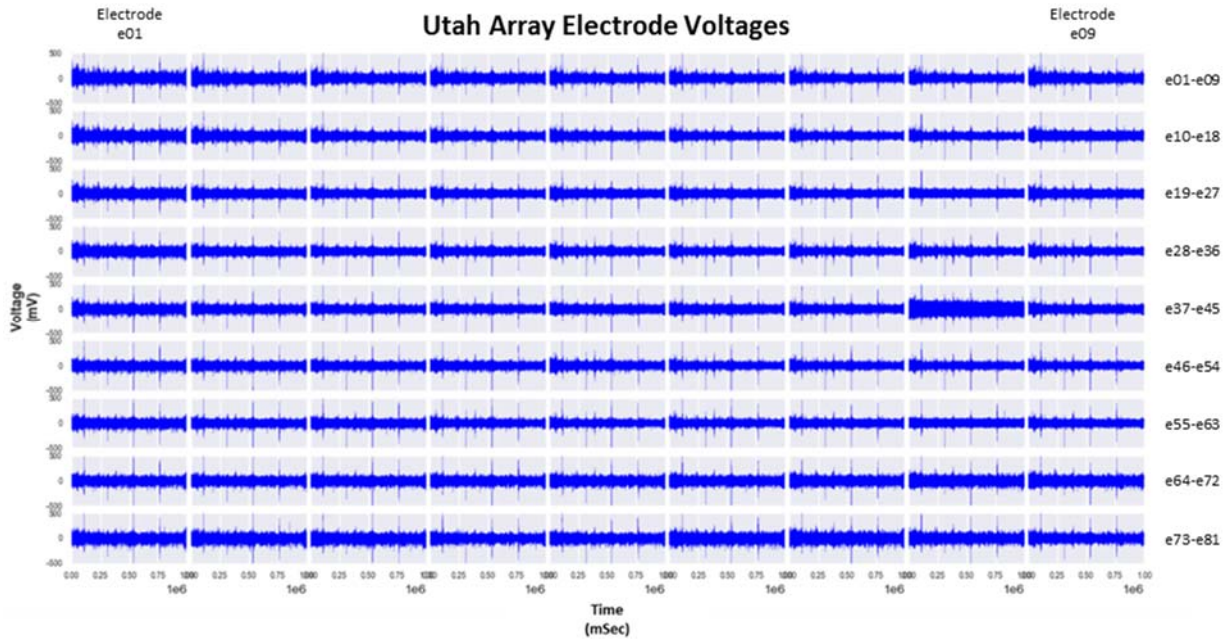

Figure 5 – 81 time series voltage plots of the Utah Array electrodes across 1,000-seconds, each plot.

In this figure, electrode 1 is located in the upper left while electrode 81 is at the lower right so as to form a 9x9 matrix from left-to-right and top-to-bottom. The electrode signals here all show similar characteristics as the prior figure, showing randomized peaks that sometimes dominate for very short moments. More importantly, these time series plots all show a mean of 0 mV and are therefore each a Gaussian distribution of data points. For Figure 5, it is worth observing that electrode 44 (row 5, column 8) generally shows higher voltage inputs on average than any of the other electrodes.

COMMON AVERAGE REFERRANCE

Although the provided dataset was initially voltage referenced to a subdural / epidural surface, an effort was made here to investigate if a Common Average Reference (CAR) application upon the provided Utah Array data would be of any further benefit to noise filtering of the electrode signals [9]. In so doing, the voltage at any one electrode is subtracted relative to the mean of a smaller subset of the Utah Array electrodes, specified as the reference voltage, V_{REF}. The CAR is computed as follows:

$$V_i^{CAR} = V_i^{ER} - \frac{1}{n}\sum_{j=1}^{n} V_j^{ER} \quad (1)$$

Where,

$$V_i^{ER} = V_i - V_{REF} \tag{2}$$

In this way, this minor investigation attempted to average electrodes e18, e26, e28, e44, e62, and e78, the selection of which are resultant from a cross-correlation analysis of electrode pairs with respect to e01, as shown in Figure 6, below. Intentionally, none of these electrodes possess strong cross-correlations, but rather, share similarly low auto-correlation values. As a basis to compare, e44 was chosen for its amplified voltage inputs (Figure 5) as well as its low scoring auto-correlation result (Figure 6). Unfortunately, for many of the evaluations forthcoming in this paper, none of these analysis showed valuable insights using CAR, and thus, no further effort was expended in this direction. However, an improved analysis of the CAR application could possibly result by improved electrode choosing for V_{REF}, by including only the smaller auto-correlation valued electrodes as compared to what was selected here.

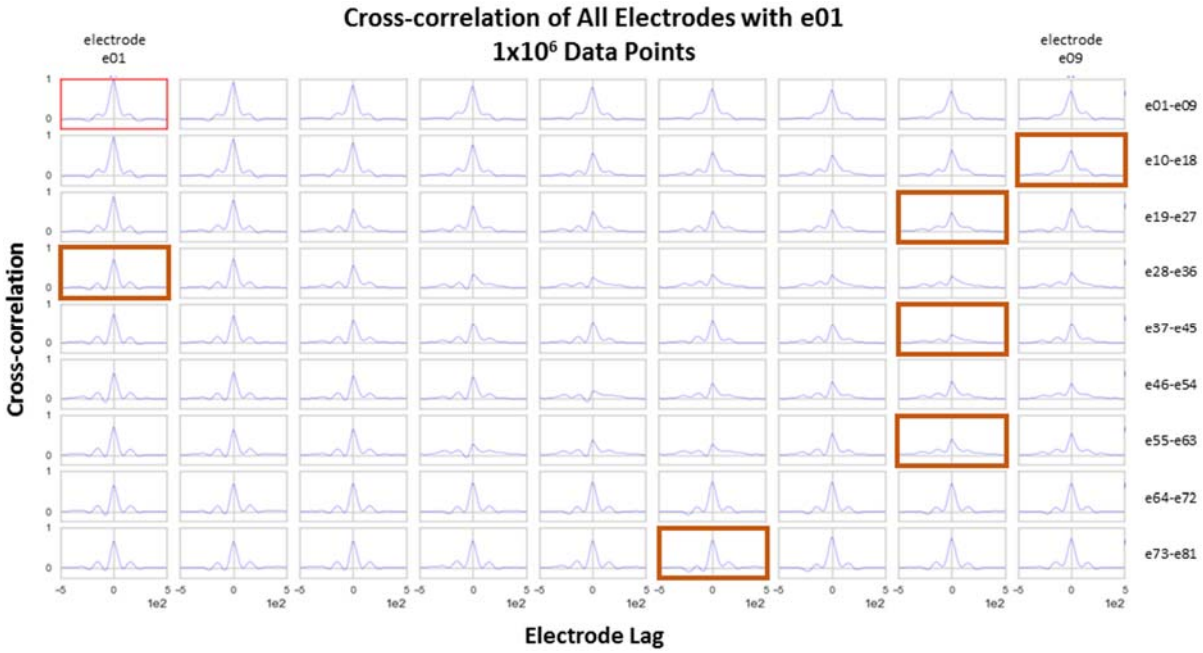

Figure 6 – Cross-correlation plots of the Utah Electrode Array, measured against electrode e01 at 1-million data points.

HEATMAP PLOTS

To further build insight to the basic time series plots given in Figure 4 and Figure 5, a heatmap plot is employed to simultaneously visualize each electrodes voltage inputs at any

one specified time sample via voltage color encoding. The associated heatmap color scale contains the spread of both negative and positive Utah Array voltages that consequently map to a range of color shades between blue and red, respectively (Figure 7). Physically, the Utah Array arrangement of electrodes is equivalent to the heatmap positions, and also to the arrangement shown in Figure 5, a 9x9 matrix (from e00 to e80). The mathematical mapping of each voltage value to the prescribed color scale is henceforth accomplished via a Look Up Table (LUT), of which is discussed in the next section.

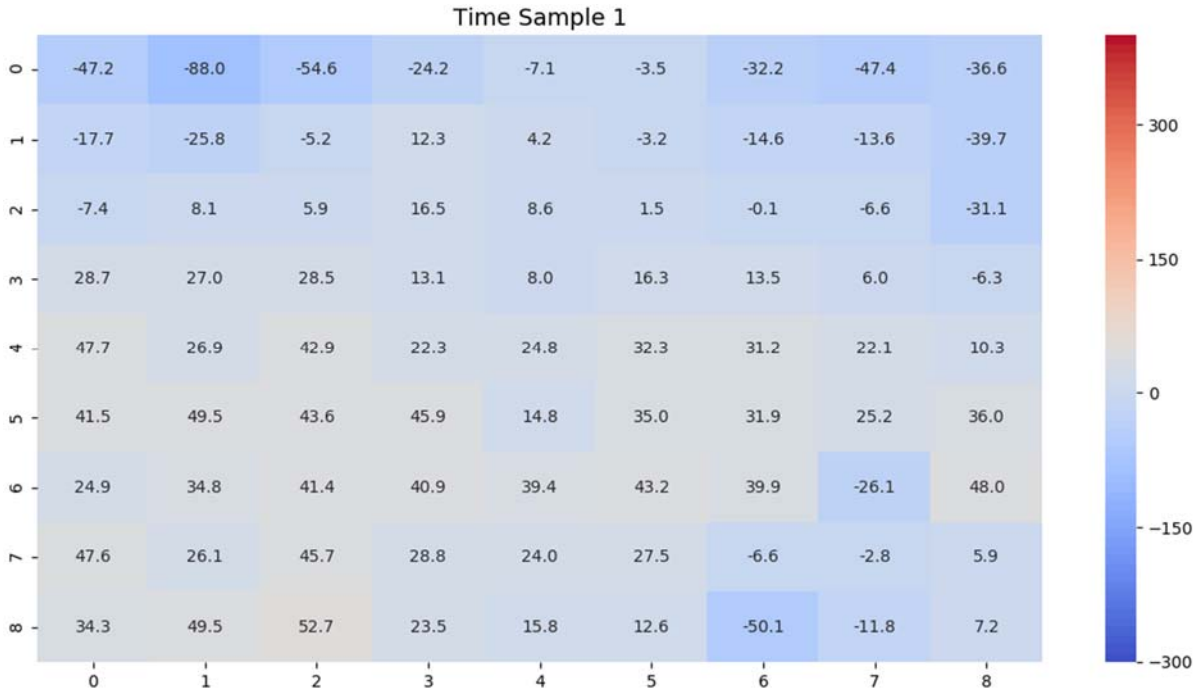

Figure 7 – A heatmap plot of a single time sample of the Utah Electrode Array voltages that map to a color scale (units in mV).

For purposes of visualizing a specified time window of Utah Array activity via a heatmap, Figure 8 demonstrates a 100-mSec window of elapsed time for all 81-electrodes concurrently. Indeed, the depicted time series shows a very dynamic environment of fluctuating voltages. In this case, the boundaries between intra-cortical activities of electrodes lacks contrast due to the chosen color scale, a linear LUT. Even so, when the time series expands to a 400-mSec window, as shown in Figure 9, a larger more dynamic range of activity is consequently observed.

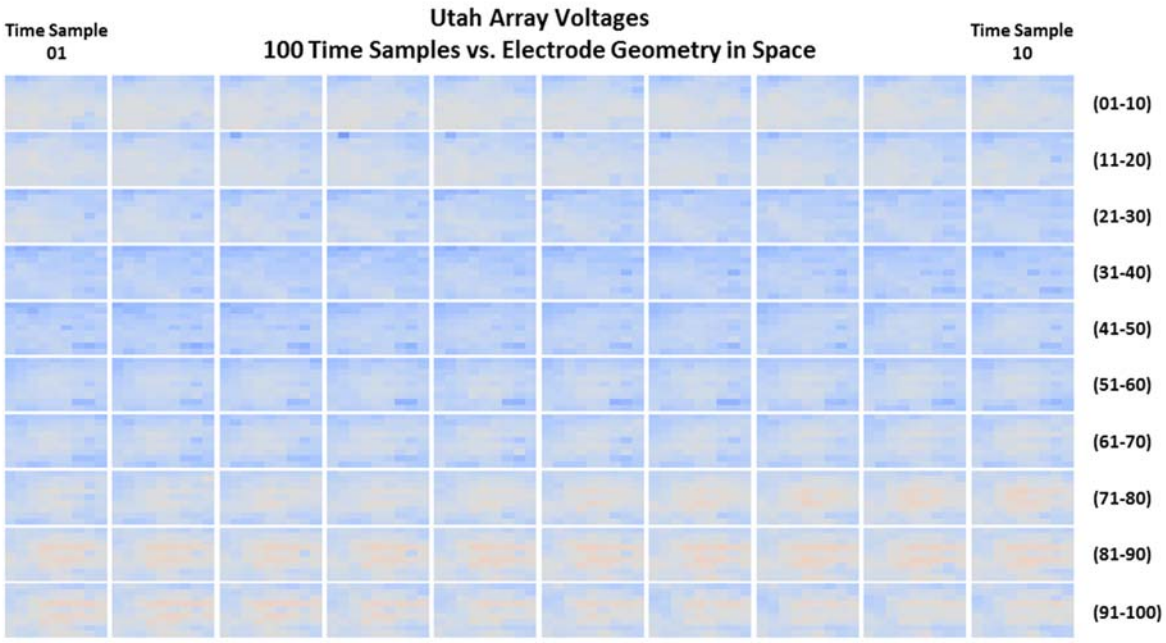

Figure 8 – A 10x10 series of heatmap plots depicting 100-mSec of Utah Array voltages for all 81 electrodes, each plot.

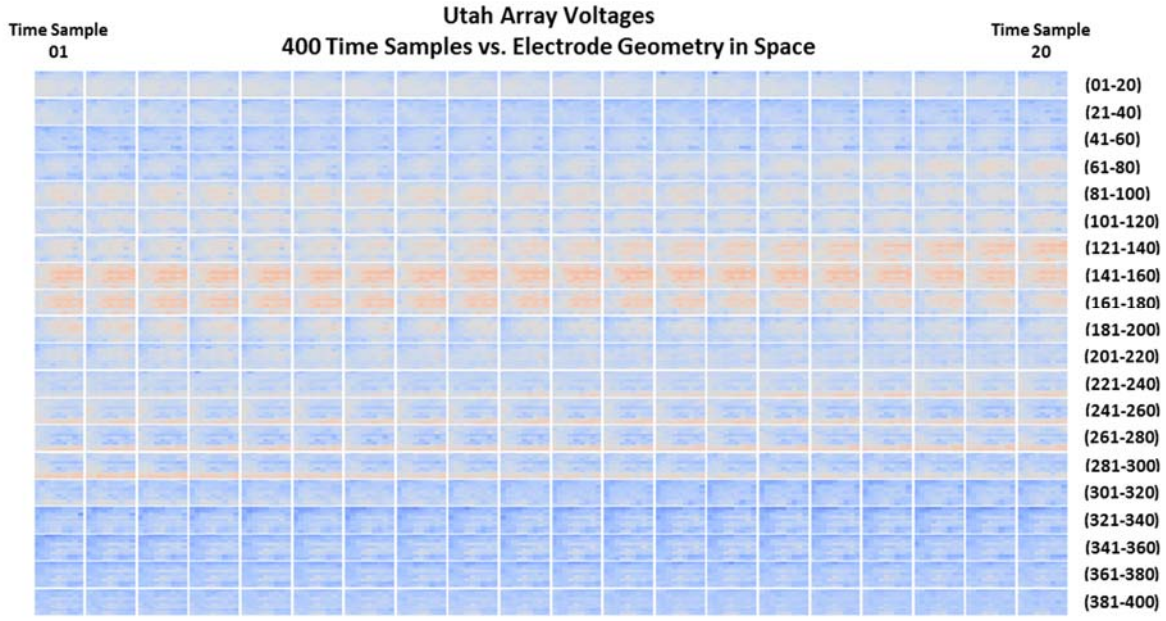

Figure 9 – A 20x20 series of heatmap plots depicting 400-mSec of Utah Array voltages for all 81 electrodes, each plot.

LOOK UP TABLE (LUT) TRANSFORMATIONS

The objective to modifying electrode voltages onto a heatmap plot via a modified transformation function and Look Up Table (LUT) is to better visualize the wide range of each electrodes signals without over-saturating the plot result. Because the dataset used in this paper contains electrode voltages that over time are nearer the origin than the minimum or

maximum of their dynamic voltage ranges (see Figure 4, Figure 5), plotting electrode voltage data with modified LUT's maps these nearer-to-zero voltages to a broader range of the color scale [10]. For example, given a heatmap voltage plot with color scale limits from ± 100 µV and dataset ranges from ± 500 µV, an electrode voltage exceeding these color scale limits is mapped to either the minimum or maximum of that color scale range. In this scenario, the heatmap image loses granularity and becomes less interpretable. Therefore, a LUT helps to modify this type of detrimental outcome. In the case of a color pixel, consisting of red, green, and blue (RGB) attributes, a LUT modifies each of these according to the applied transformation function. A clear example of modifying pixel values, for both linear and non-linear functions, are given in Figure 10 below. In this figure, the mathematically defined transfer function (both the dashed line and the pink line) establishes the LUT mappings by concluding new RGB values at the y-axis [11]. When compared to a linear LUT (Figure 10, left), the non-linear LUT (Figure 10, right) shows the re-mapped RGB values of larger magnitude. This type of result greatly improves image contrast, as the differences in magnitude between the updated LUT and the pre-existing LUT are now increased. Thus, applying a LUT to the Utah Array voltage data is the chosen method to effect improved meaningfulness from the resulting heatmap plots employed for analysis [10].

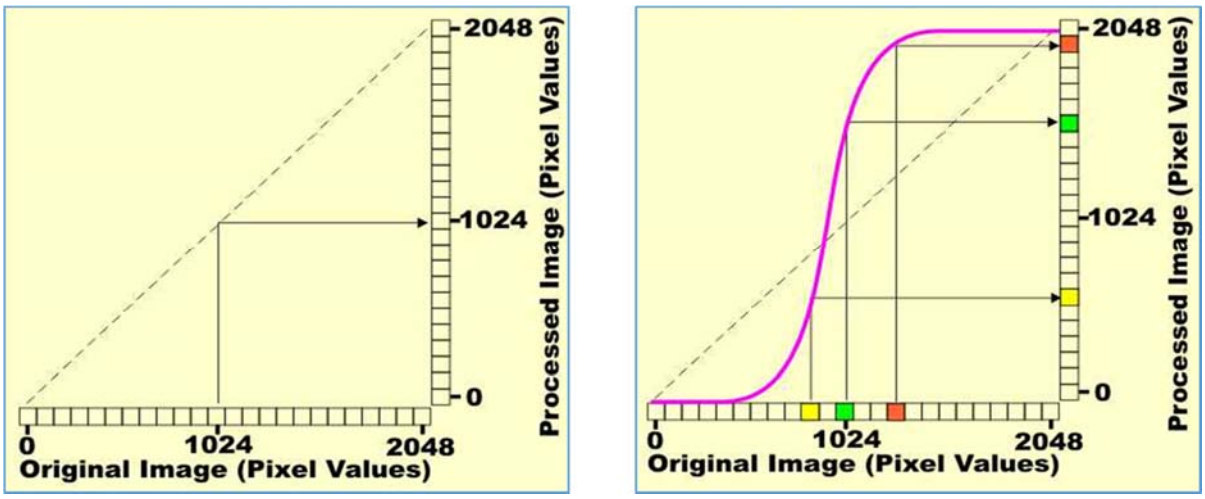

Figure 10 – Depiction of both a linear (left) and a non-linear (right) LUT applied to pixel attributes.

The four heatmap plots below in Figure 11 demonstrates such various LUT's upon a single time sample of the 9x9 Utah Array geometry, where the following transformation functions

are used to define the applied LUT: $[f(x) = x, f(x) = x^2, f(x) = x^3,$ and $f(x) = log_{10}(x)]$. Just as Figure 11 shows, each of the associated transformation functions emphasizes the varying quality of contrast, consequently tweaking the image interpretability.

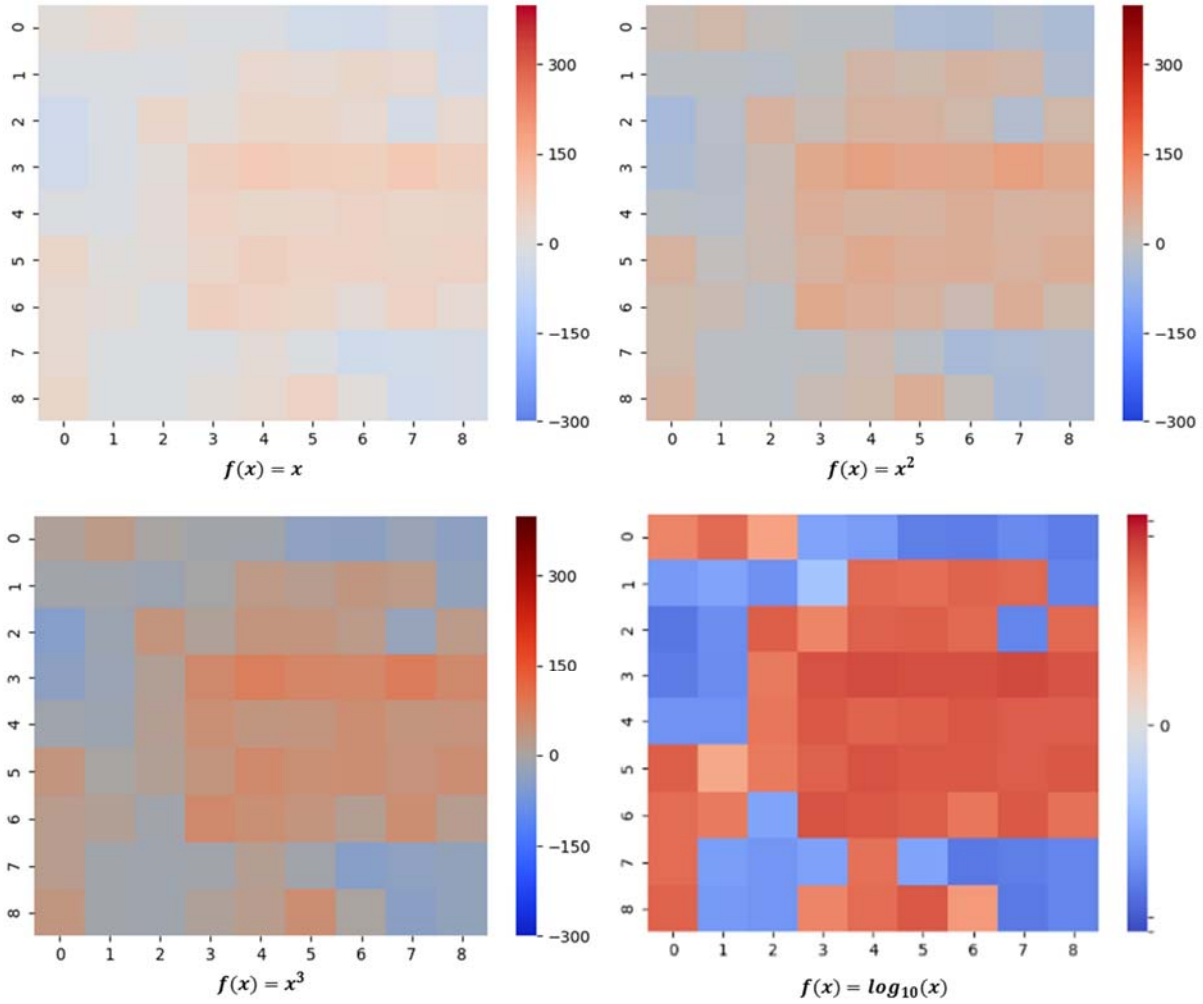

Figure 11 – Comparison of Applied Transformation (LUT) Functions. Properly implementing a LUT mitigates over-saturation of the visualized data, thus delivering more meaning to both human and machine perspectives.

To further exemplify the improved visualization outcomes of appropriately choosing a transformation function upon the Utah Array dataset, a single time sample is plotted for comparison in below Figure 12, using both a linear LUT and a logarithmic LUT, to all 81-electrodes of a heatmap plot. Clearly, the right plot makes interpretation of the values much easier to see, specifically, where the more (or less) active regions of the Utah Array are located.

Another example is given in Figure 13, where the same LUT comparison is made upon a correlation coefficient heatmap (discussed later). Observe here the improved contrast when employing the log function to the right plot. Clearly, the updated logarithmic transformation function results in improved interpretation.

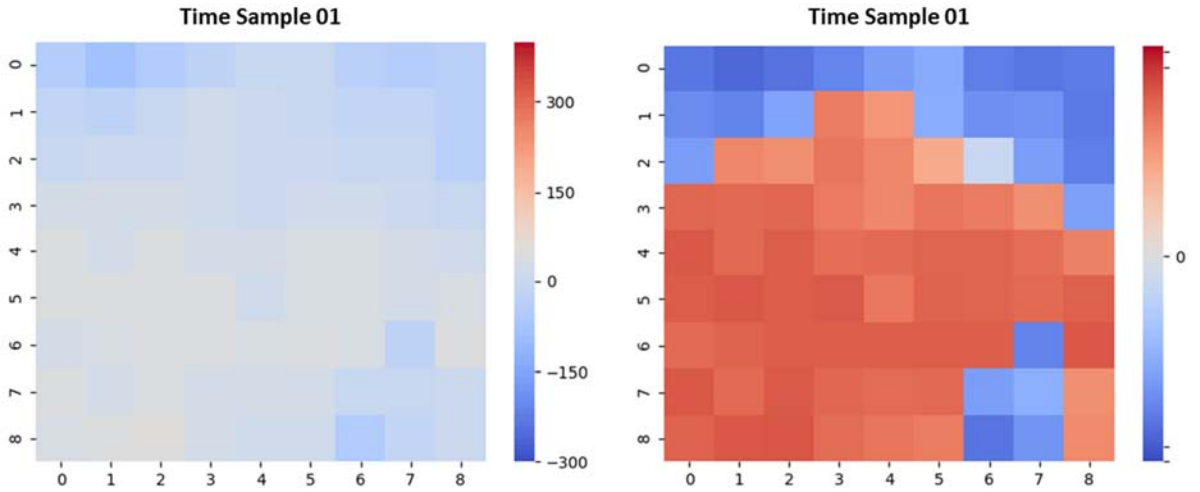

Figure 12 – Heatmap plots and their applied LUT's showing raw dataset electrode voltages across the 9x9 Utah Array for each of the 81 electrodes for improved image contrast: linear LUT (left), logarithmic LUT (right); time sample 1 of the dataset.

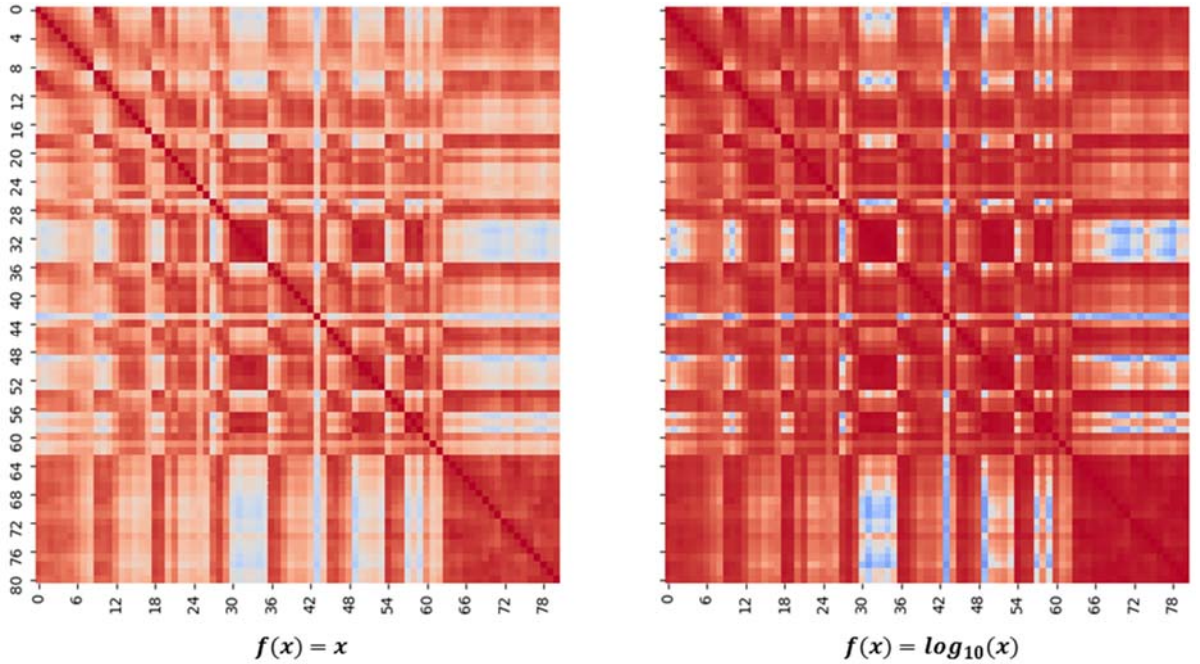

Figure 13 – Comparison of correlation coefficient heatmap plots across 1000-samples (raw data) after applying a specified Look Up Table (LUT) transformation such that the plot does not oversaturate.

TRADITIONAL ANALYSIS

A traditional analysis of an Utah Electrode Array dataset pertains mostly to techniques using statistics. In order to apply such methods, it is necessary to first demonstrate that the dataset of electrode values holds to a normal Gaussian distribution. Such an outcome is previously shown in Figure 5, where mean values for each of the electrodes average toward zero across the time series of 1×10^6 plotted data points. Therefore, statistical investigations using moving averages and moving standard deviations, histograms, and probability densities are at our disposal. Equally, scatterplots and joint plots both compliment the aforementioned and provide further insight to the underlying data. In this paper, both scatterplots and joint plots are important tools giving quick visualization results toward electrode interdependencies. Another visualization tool closely related to the scatterplot is the correlation matrix, where electrode interdependencies are shown as a subset of stacked electrode scatterplots within a grid. This method consequently results in a matrix diagonal of self-identity, that is, the normal distribution for that electrodes data is shown in-place of the scatterplots auto-correlation result.

Yet another approach to identification of electrode voltage patterns from the Utah Electrode Array dataset is to investigate via a signal analysis approach called cross-correlation. This technique evaluates pair-wise the various combinations of electrode signals, as the technique is used for inspecting signal similarity comparison. A signal that is correlated to itself is said to be auto-correlated, and possesses a magnitude of 1. In the latter parts of this section, explanation is given not only to cross-correlation results such as this, but also toward the identification of specific lag times containing the suspect signal similarities between the compared electrodes. For example, if at an initial time a signal is received upon an electrode, it is of interest to identify that same signal if it is again received upon another electrode at a specified time delay called lag. This repetitious behavior is observed by way of cross-correlation plot results.

MOVING AVERAGES AND MOVING STANDARD DEVIATIONS

The fundamental importance of a moving average calculation is to smooth a time series of data for improved interpretability [12]. This is especially so for an electrode signal, where its characteristics become key features in Utah Electrode Array time series visualizations. In the specific case of this study, a central moving average (CMA) is employed against each of

the 81 electrodes across the first 10-seconds of data points. Figure 14 below shows such a comparison between the raw signal (top) and the moving average signal values (middle) for electrode e01. The figure clearly shows an improved time series visualization for overall interpretability. For this investigation a bin size of $n = 10$ was used to calculate the moving average of mean voltages, where mathematically this is described by:

$$\bar{v}_{CMA} = \frac{v_M + v_{M-1} + \cdots + v_{M-(n-1)}}{n} \tag{3}$$

$$= \frac{1}{n}\sum_{i=0}^{n-1} v_{M-i} \tag{4}$$

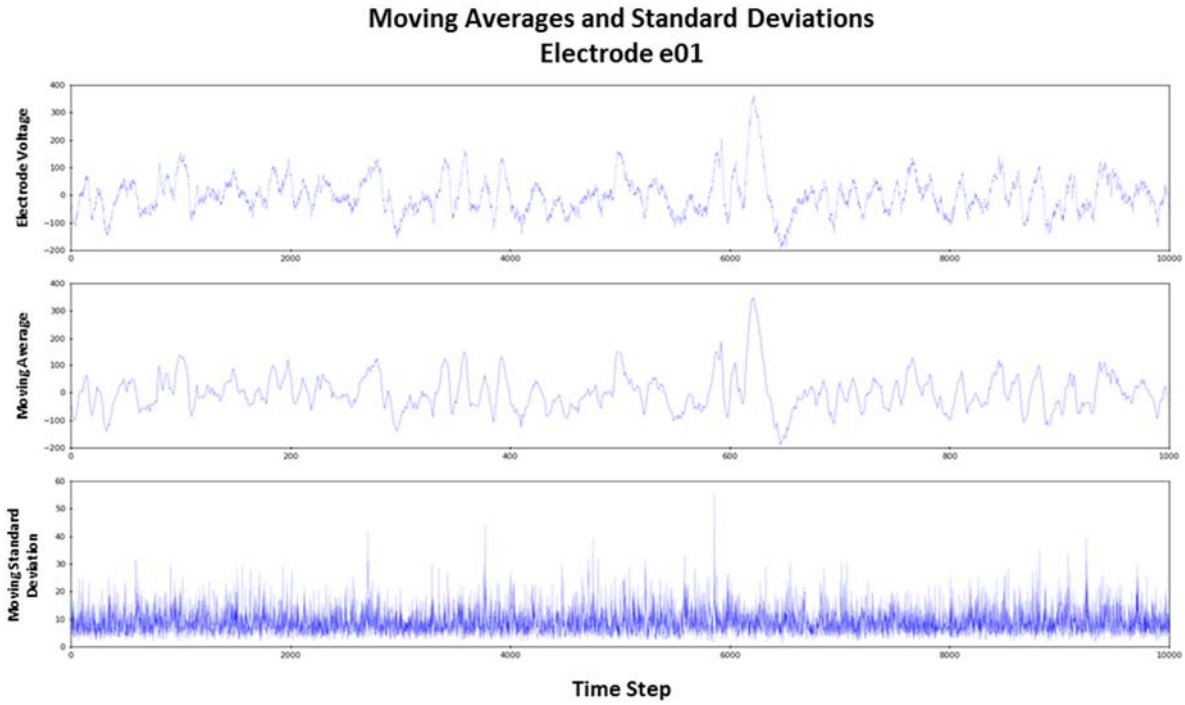

Figure 14 – Time series plots for electrode e01 depicting: raw electrode voltages (top), moving averages (middle), and moving standard deviations (bottom).

For the CMA, all 81 electrodes when plotted in this fashion resulted in similar time series plot characteristics without noticeable differences, except for electrode e44, where the moving average persists as a noisier signal when evaluated against the others. In comparison, the moving standard deviations such as shown in Figure 14 (bottom) for electrode e01 depict a mean of ≤ 10. This remains true for the majority of the 81 electrodes with a few exceptions, where the mean is ≥ 15: e18, e26, e44, and e62 (see Figure 15). The purpose for

establishing the standard deviation of a signal is to indicate the amount of variation between data points. This is mathematically expressed as [13]:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(v_i - \bar{v})^2}{n-1}} \qquad (5)$$

In the context of this paper, moving standard deviations are taken across $n = 10$ data points.

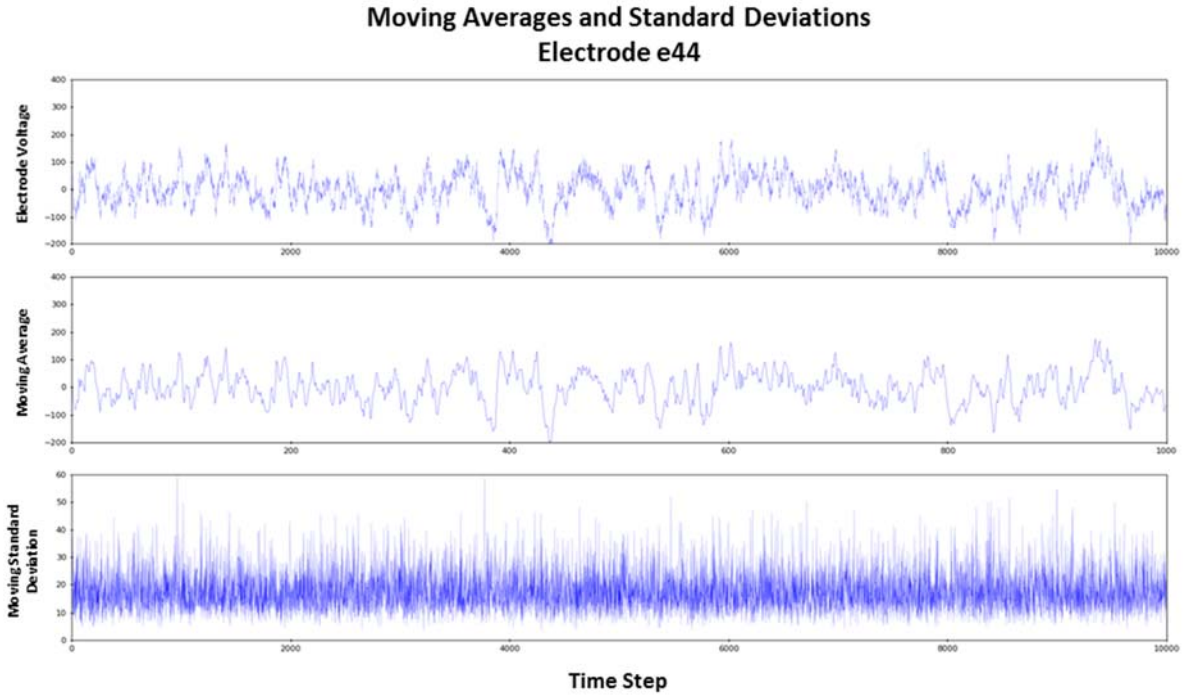

Figure 15 – Time series plots for electrode e44 depicting: raw electrode voltages (top), moving averages (middle), and moving standard deviations (bottom). Observe the higher moving standard deviations as compared to electrode e01, above in Figure 14.

HISTOGRAMS, NORMAL DISTRIBUTIONS, PROBABILITY DENSITIES

A histogram plot is one of the more important visualization techniques in statistics used to quickly inspect various characteristics of a continuous variable dataset, such as its distribution, skewness, or outliers [14]. For this plot type, the distribution of data is evaluated and arranged into bins, where the area of each bin conveys the frequency of occurrence for those binned dataset voltages. Figure 16 below demonstrates this construct for electrode e01 using several bin sizes (10, 100, and 1000) and 1×10^5 data points, where each of the distributions show increasingly improved plot resolutions. The results clearly show that they are non-skewed, nor do outliers exist for this subset of Utah Electrode Array data. In this

same manner, each of the 81 Utah Array electrodes was shown to possess very similar characteristics as those shown in Figure 16 – a normal Gaussian distribution.

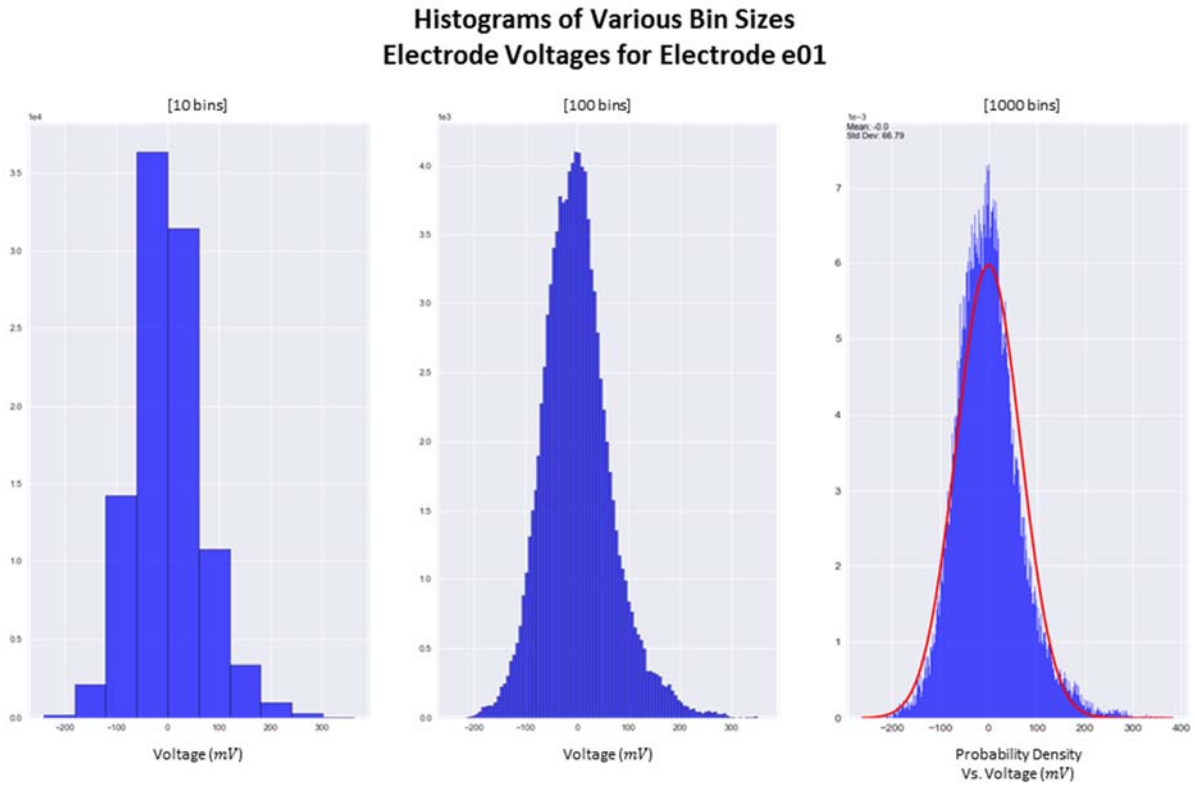

Figure 16 – Various histogram bin sizes (left to right: 10, 100, 1000) showing typical characteristics of a normal Gaussian distribution of electrode e01 data.

For the right-most plot, an additional element is shown that mimic's the histograms shape, called the probability density function (PDF). This curve is described by the relation [15]:

$$PR(x \in A) = \int_A \rho(x)dx = 1 \qquad (6)$$

where the probability, PR, that random variable x resides within dataset A is computed as an integral of the density function $\rho(x)$, with the integration result, the area under the curve, is equal to 1. For a Gaussian distribution, the density function is:

$$\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad (7)$$

Consequently, the resultant PDF curve gives a method with which to compare each of the 81 electrode signals in a standardized manner, such as is shown in Figure 17, below. Plotting

each of the electrode density curves shows that all 81 signals are expected to behave similarly without skewness. In actuality, this figure was derived using a Kernel Density Estimation (KDE) function using a Gaussian kernel, such that a non-parametric estimation of the desired PDF is achieved.

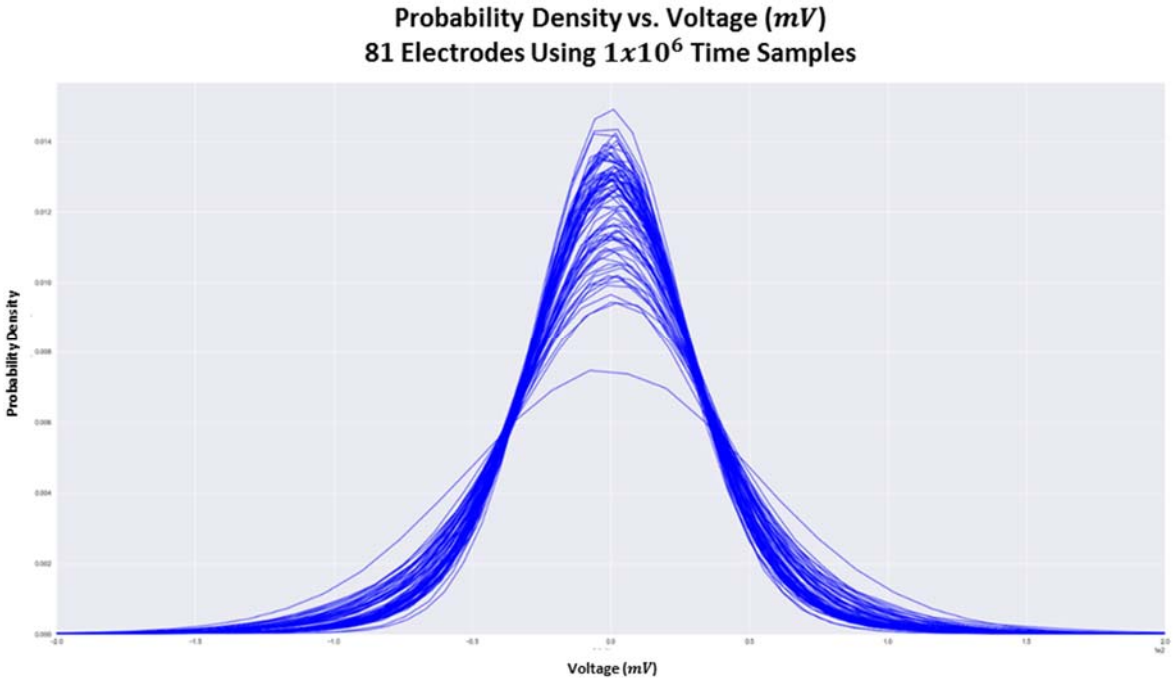

Figure 17 – Probability density curve estimates for all 81 electrodes using a Kernel Density Estimation (KDE) and a Gaussian kernel.

Another manner with which to utilize a histogram is to capture each of the Utah Array electrode values at a single specified time. As opposed to Figure 16, where a single electrode was evaluated for many time samples, this new approach permits observation of the moment-to-moment flux of both the PDF and histogram attributes of the Utah Electrode Array. For illustration, both Figure 18 and Figure 19 give such results for the time series plots of 100 and 400 Utah Array samples respectively. In this fashion, the PDF provides interpretation of the variance in both standard deviation and a histograms mean for the given voltages, where particular time samples indicate if a specific voltage range (bins) becomes dominant over others. When assessed alongside the PDF in this way, the likelihood of improved feature identification for neural pattern recognition consequently improves. The next section of this paper introduces scatterplots and joint plots as another approach to more deeply understanding the provided datasets contents.

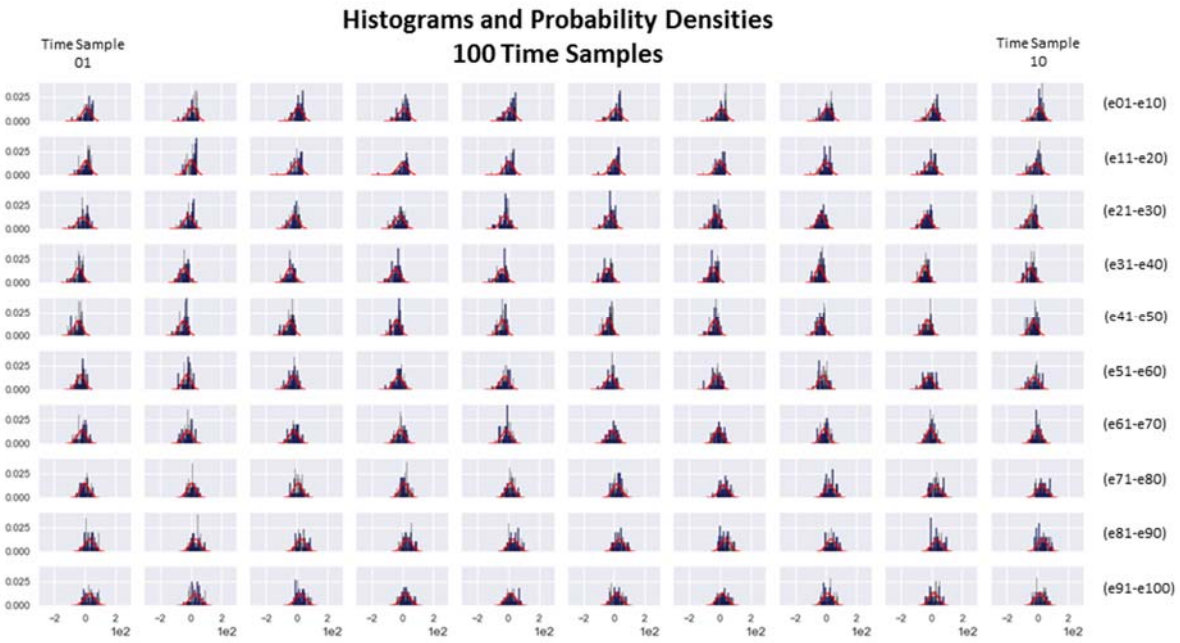

Figure 18 – 100 time series histogram plots visualizing all 81 electrodes of the Utah Array at a single instance in time, each.

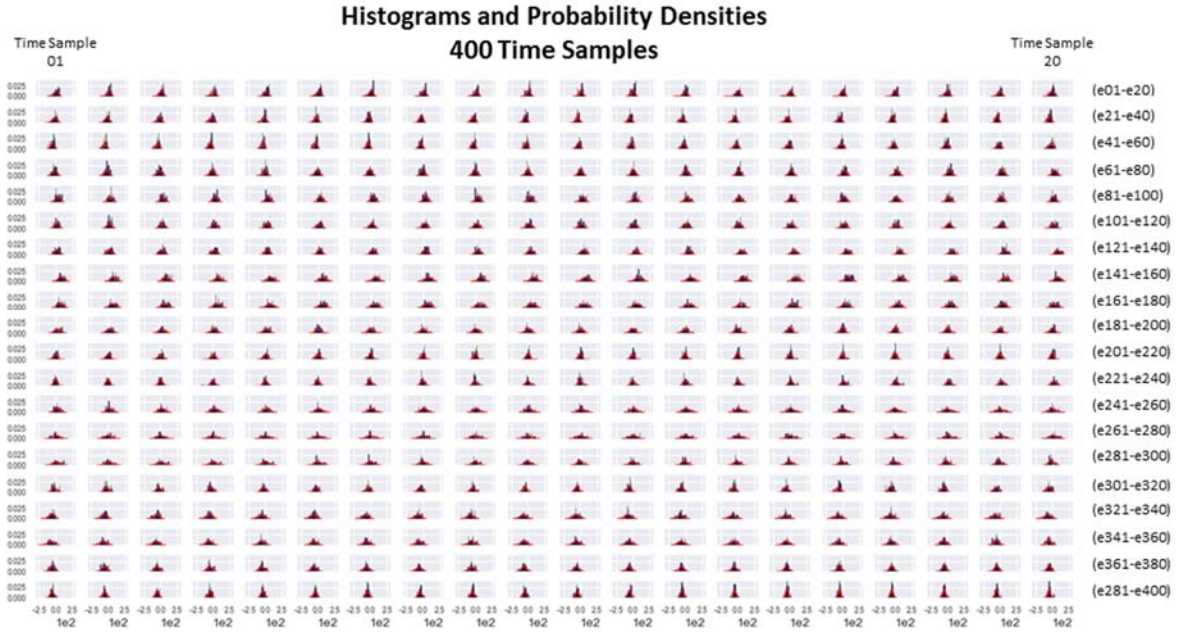

Figure 19 – 400 time series histogram plots visualizing all 81 electrodes of the Utah Array at a single instance in time, each.

SCATTER PLOTS AND JOINT PLOTS

A scatter diagram, or scatter plot, associates continuous variables x and y, dependent and explanatory, in a simple to interpret geometric result that offers important statistical analysis capabilities upon a given pair of dataset electrodes [16] [17]. Such plots are one of the best methods available to quickly evaluate electrode dependencies in 2-dimensions for

direct comparison and interpretation. As such, meaningful insights are garnered using methods of correlation, linearity, outliers, and clustering [17]. As one of the more important of these tools, correlation - a measure of the strength of the relationship between variables [18], is demonstrated below in Figure 20, where scatter plots of electrodes e01 and e81 are given for comparison. For an auto-correlation (left plot), the plots x and y -axis each use the same data points (electrode e01 here) that also results always in a correlation magnitude of 1 (versus a non-correlated outcome of -1). In the specialized case of an auto-correlation, the plotted voltages map to both x and y variables in the same way, resulting in a linearly positive slope of 45°. In the opposing plot, a more diffuse scatter plot of electrode voltages is shown and consequently an observed correlation magnitude of < 1 is verified with the regression line (in red). If the results plotted were instead to have a negative correlation, then this regression line would also have a negative slope.

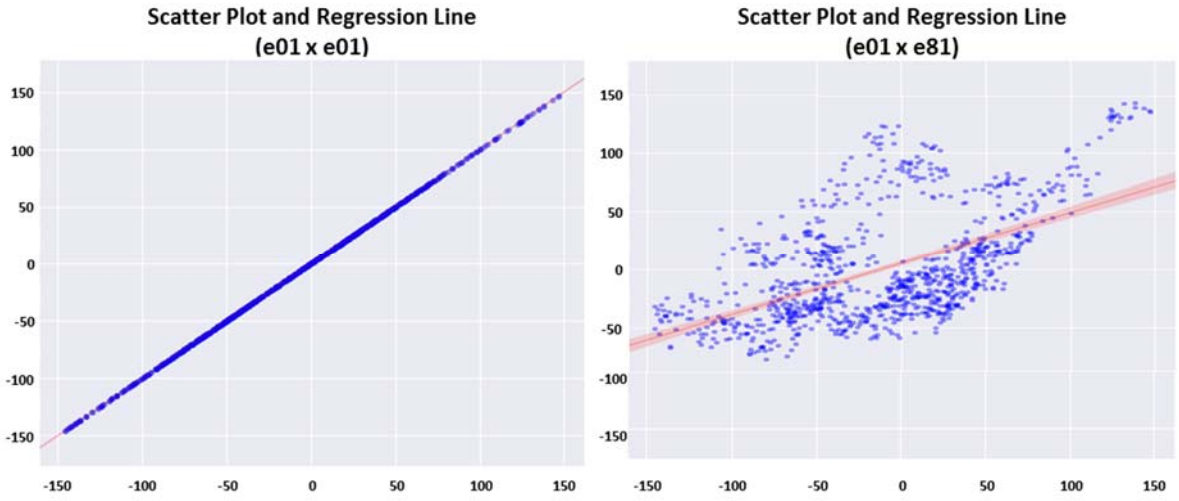

Figure 20 – Scatter plots depicting correlation results and regression line: an auto-correlation of electrode e01 (left), and the scatter plot correlation of electrodes e01 and e81 (right).

So as to mathematically associate the regression line and the correlation coefficient, r:

$$\mathrm{r} = \mathrm{b}_{\mathrm{Y}\cdot\mathrm{X}}\left(\frac{\mathrm{S}_{\mathrm{X}}}{\mathrm{S}_{\mathrm{Y}}}\right) = \mathrm{b}_{\mathrm{X}\cdot\mathrm{Y}}\left(\frac{\mathrm{S}_{\mathrm{Y}}}{\mathrm{S}_{\mathrm{X}}}\right) \qquad (8)$$

where $b_{Y\cdot X}$ and $b_{X\cdot Y}$ are each the slopes of the regression line for predicting Y from X and X from Y, respectively. Correlation then, is a function of the slope of the regression line and also of the variables ratio of standard deviations, where this ratio acts as the agent of transformation from a regression slope into a standardized correlation slope [16].

Mathematically, as r decreases from the $r = +1$ auto-correlation result (left plot) toward zero, the plotted data changes from the prescribed linear slope of 45° into an ellipsoid (right plot). Continuing on in this way, when $r = 0$ the plotted data resembles a circular shape [17] and the regression line has a slope of 0°. The ultimate use of a regression function is to seek a response from the explanatory variable. It is important to note that although correlation is symmetric between its variables, this is not true for regression.

Another statistics tool used for visualizing the combined outputs of both a scatter plot and a Kernel Density Estimation (KDE) function is called a joint plot. As before, this plot type is intended to improve human interpretations of variable dependencies such as those found in the Utah Electrode Array dataset [17]. In comparison to Figure 20, the primary plot area of Figure 21 uses isotropic contour lines with shading to indicate variations in voltage densities. Additionally, the related KDE functions are projected onto the opposing plot axis, permitting a more clear visualization of the dependent and explanatory variable distributions. For the figure below, the left plot shows the auto-correlation of e01 where the right plot shows a more diffused, elliptically shaped, voltage clustering using electrodes e01 and e81. These joint plots show the usefulness of synthesizing multiple aspects of a density distribution, and by additionally including the pearson-R and probability scores, p, onto the plot, the potential discovery of the less visible relationships consequently increases. Note that the pearson-R value is equivalent to the regression slope as previously discussed above, and is consequently related to the correlation of the compared electrodes. In the next section of this paper, the constructs previously discussed regarding scatter plots and correlations are combined to further establish improved insights into the Utah Electrode Array dataset attributes.

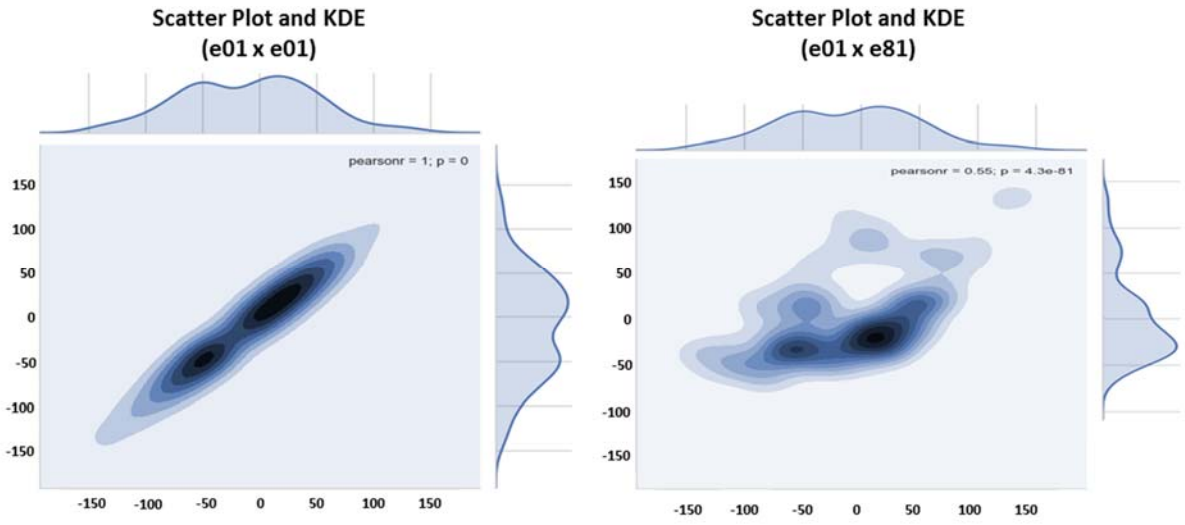

Figure 21 – Scatter plots depicting correlation results and KDE density distributions: an auto-correlation of electrode e01 (left), and the scatter plot correlation of electrodes e01 and e81 (right).

COVARIANCE (CORRELATION) MATRICES

A covariance matrix permits data visualization assessments at higher-levels than is offered from an individual scatter plot observation, where in a correlation matrix, an increased number of scatter plots are simultaneously compared. In fact, this plot type is well suited to investigate the Utah Electrode Array dataset, as it is often used in studies of multivariate statistics, where spatial distances allow the mutual relationships between variables to more easily be exposed [19]. Figure 22 demonstrates this capability nicely, as scatter plots of electrodes e01 through e09 (a single row of the Utah Electrode Array) are constructed according to the correlation matrix layout. To understand this, begin with the left-most column (electrode e01), where the upper scatter plot is an auto-correlation of e01, just as was shown in Figure 20.

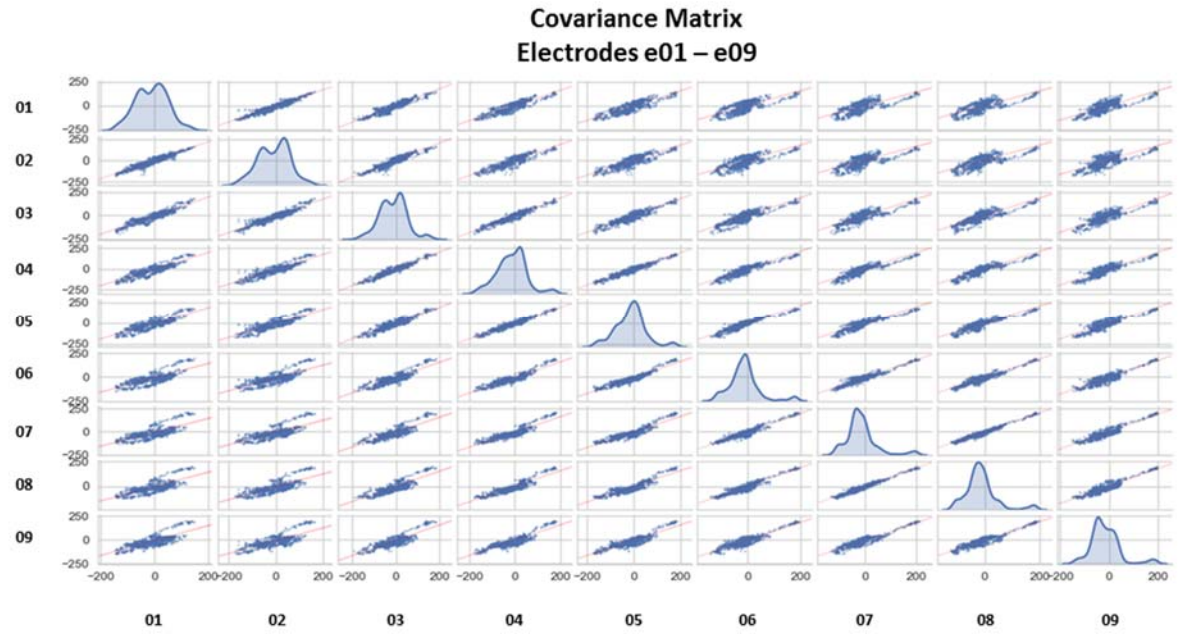

Figure 22 – Covariance (correlation) matrix of electrodes e01 through e09, showing positive correlations and KDE's as diagonal auto-correlation results.

However, for this particular result, the auto-correlation plot is replaced with a Kernel Density Estimation (KDE) as a more meaningful result, where this same output is employed for all matrix auto-correlations shown along the correlation matrix diagonal. Thereafter, moving down the column, electrode e01 is compared against the next row (electrode e02), and so on until electrode e09 is compared against e01. In the same manner, this column-by-column process is interpreted for all consecutive columns. Furthermore, as a consequence to the non-directional correlation coefficient (discussed later) where,

$$r_{XY} = r_{YX}, \tag{9}$$

the matrix plots are nearly symmetric about its diagonal. The subtle differences are mathematically given as a simple sign reversal, such that maximally only half of the n scatter plots require computation, and only N plots are necessarily significant [17]:

$$\mathrm{N} = n(n-1)/2. \tag{10}$$

Building upon these constructs, where Figure 22 gives 9 electrode comparisons, Figure 23 depicts the correlation matrix for all 81 Utah Array electrodes. The result is that the diagonal is both retained to reinforce matrix symmetries, and the auto-correlation is replaced by the KDE, and on closer inspection, none of the relationships show negative pearson-R correlation values. More importantly, Figure 23 shows visible variations between scatter plots via

shading attributes, where less surrounding white space within the matrix indicates a less correlated plot.

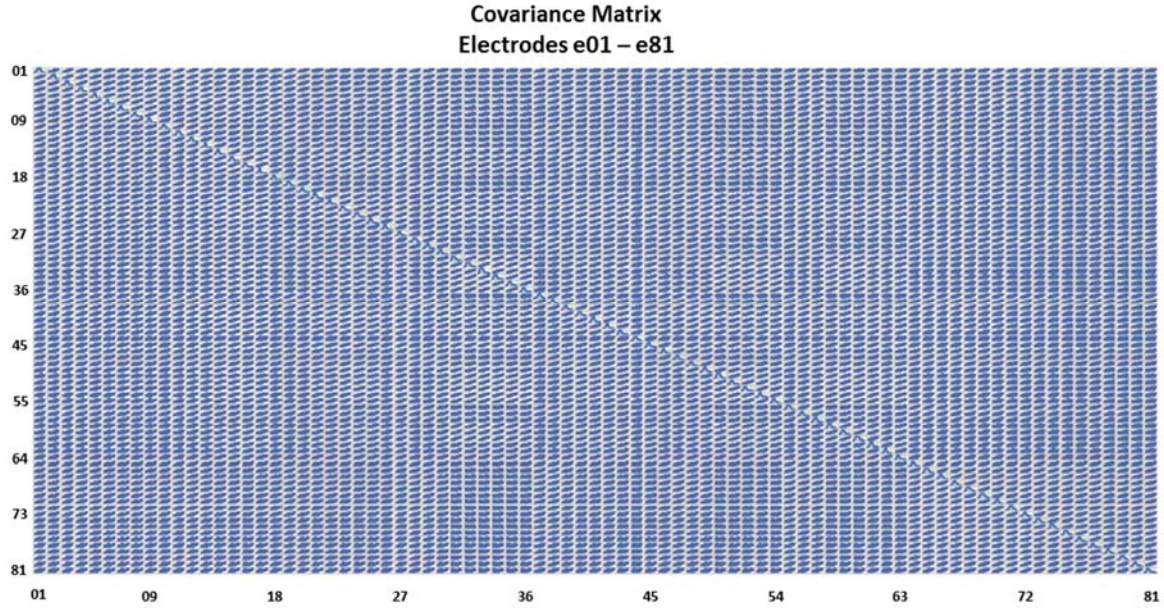

Figure 23 – Correlation matrix of electrodes e01 through e81, showing positive correlations and KDE's as diagonal auto-correlation results.

CROSS-CORRELATIONS

Up to this point, the derived Utah Array comparisons relate to statistical measures, where in particular, scatter plots are exceptionally useful and allow for direct visual comparisons of signal pairs that are spatially oriented, as is especially true for the correlation matrix shown above. However, an equivalent but more exacting tool is available, called the cross-correlation function. In this method, 2 signals $f(t)$ and $g(t)$ are compared for similarity by sliding $g(t)$ along the x-axis of $f(t)$, by a measure of time τ [20]. At each time step, the signals are multiplied together and integrated across time, as described by:

$$(f \star g)(\tau) \triangleq \int_{-\infty}^{\infty} \overline{f(t)} g(t+\tau) dt. \tag{11}$$

Figure 24 below shows an example of a cross-correlation function by way of using electrode e01 as the basis, where e01 $= f(t) = g(t)$, returning the auto-correlation result as previously discussed (Figure 20, Figure 21). A cross-correlation plot is somewhat unique in its appearance, as the typical shape is a central positive peak ≤ 1 (where an auto cross-correlation $= 1$), that sharply falls away at either side to very low values at large time shifts,

$\pm\tau$ [21]. As a point of comparison, Figure 25 gives the cross-correlation result of electrodes e01 and e33. Of particular interest here, are the asymmetric second and third peaks about the y-axis.

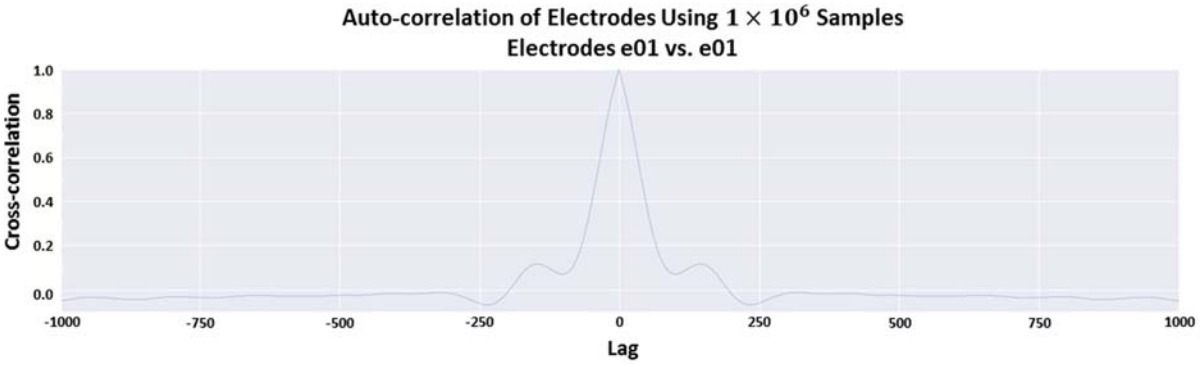

Figure 24 – The auto cross-correlation function of Utah Array electrode e01. Note the maximum peak = 1.

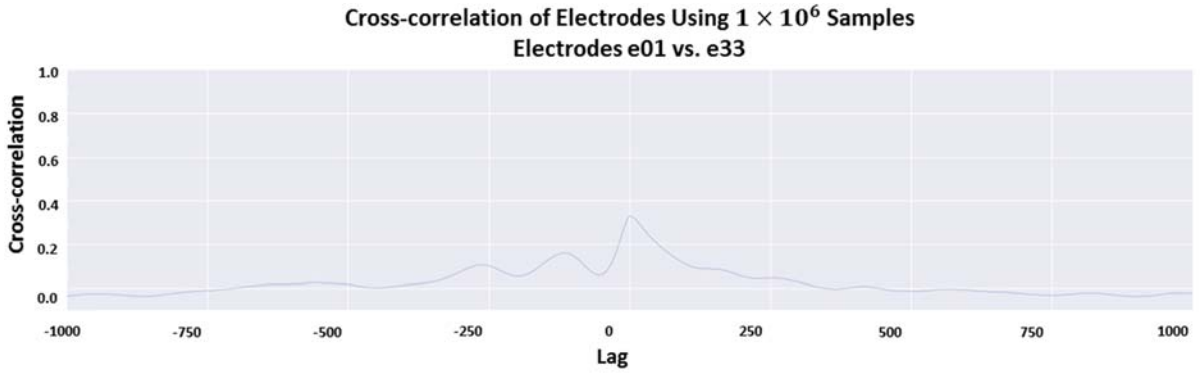

Figure 25 – The cross-correlation function of Utah Array electrodes e01 and e33. Note the asymmetric shape of the plotted function.

In Figure 26 below, electrode e01 (red box) is used as the basis for plotting the cross-correlation functions for all 1-million data points of each of the 81 Utah Array electrodes. In this way, it is easier to observe individual electrodes as more or less similar to electrode e01. In fact, many of the plots share similarities with the e01 auto cross-correlation plot, however, the more interesting characteristics pertain to those plots that are asymmetric about the central peak at $\tau = 0$. Electrodes e31 – e36, e44, and e58 – e60, as well as Figure 25 above show those asymmetries of interest. To further exemplify the construct of a cross-correlation function, Figure 27 uses electrode e33 as its basis for reversing the cross-correlation variables from (e01 x e33) to (e33 x e01). Notice here that the e01 plot result is in identically plotted (except for a sign change) as was the previously plotted electrode e33 of the prior calculation

(e01 x e33). Again, Figure 27 depicts a multitude of asymmetric plots that are of further interest to this research paper.

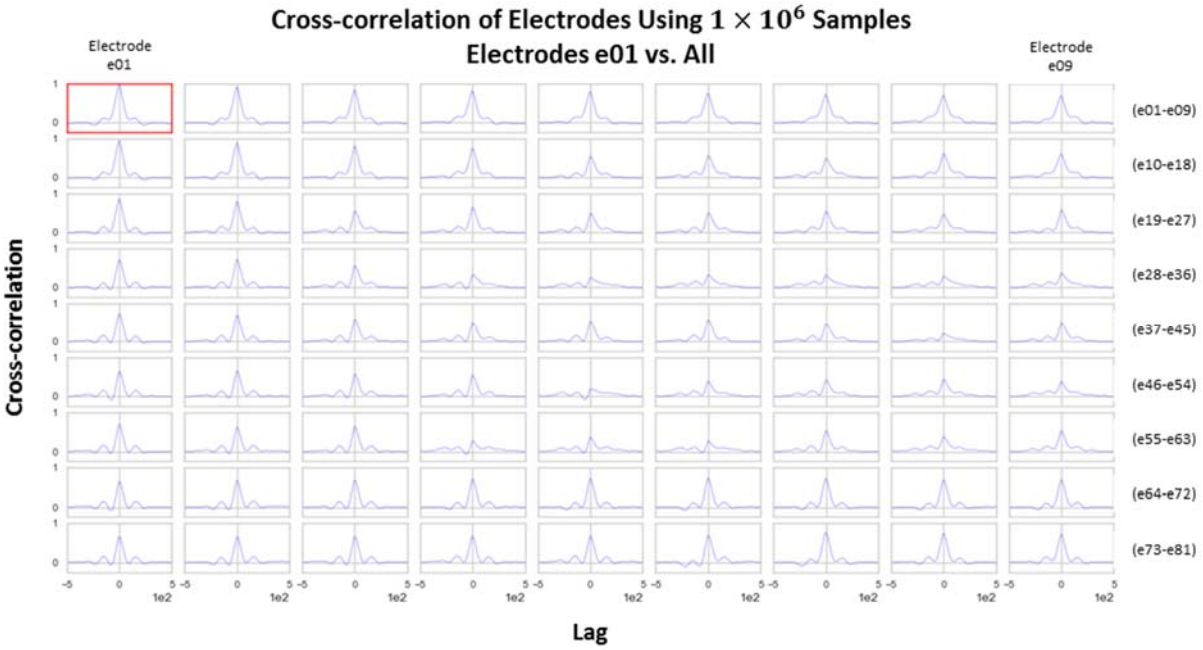

Figure 26 – Cross-correlation plots for each of the 81 Utah Array electrode against electrode e01, using 1-million data points.

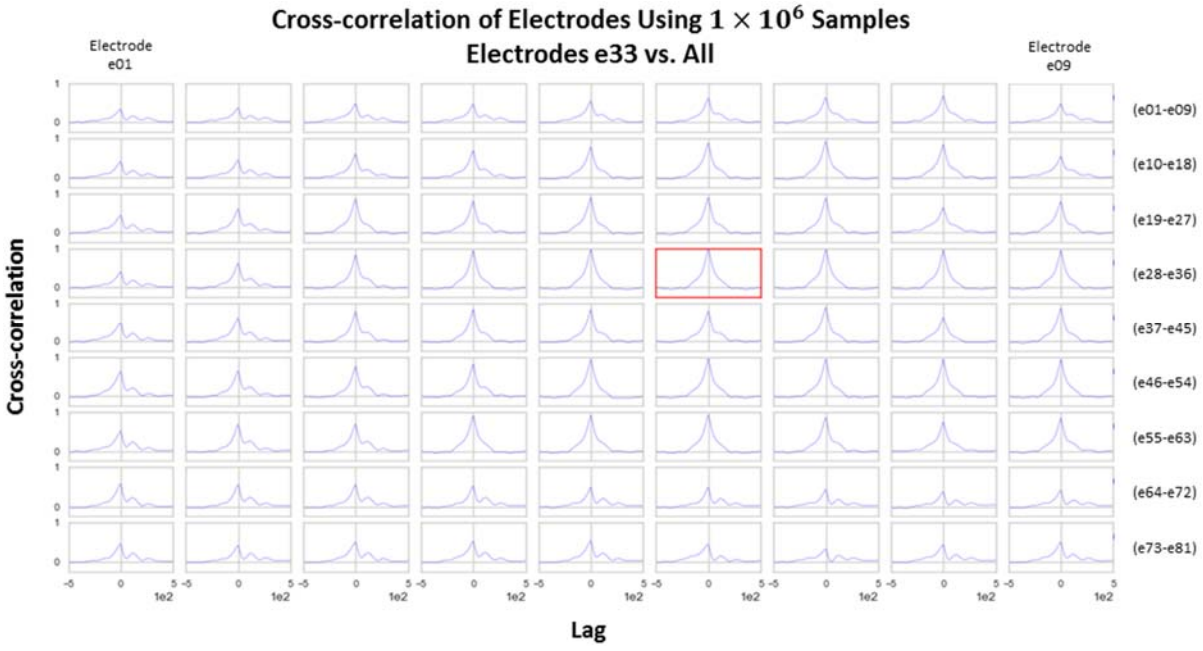

Figure 27 – Cross-correlation plots for each of the 81 Utah Array electrode against electrode e33, using 1-million data points.

CORRELATION COEFFICIENTS

The correlation coefficient, r, is a random sample estimate of both the direction and strength between variables X and Y. A positive coefficient result means that as one variable

increases, so does the other, and vice versa for a negative coefficient. However, such an outcome does not preclude that the variables are causally related, but rather that a cause does exist between them [22]. Through equation (8) above, the correlation coefficient is shown to relate to regression via the slope of the regression line and the standard deviation of the sampled variables, S_x and S_y. Another association previously discussed regarding the correlation coefficient is the correlation (covariance) matrix. By way of this matrix, the covariance (variance in X and Y together) and correlation (variance in X and Y separately) are mathematically established as:

$$r_{xy} = \frac{cov(x,y)}{\sqrt{S_x S_y}} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (12)$$

$$= \frac{cov(x,y)}{\sqrt{var(y)var(y)}}. \quad (13)$$

The correlation coefficient then, can map in the same way as the electrode covariance matrix given in Figure 23, by examining the resultant columns and rows. In this way, Figure 28 plots several histograms at different bin sizes for electrode e01, as it is the basis of all 81 electrode correlation coefficient computations. Observe that the mean of the probability density curve (KDE) is positive.

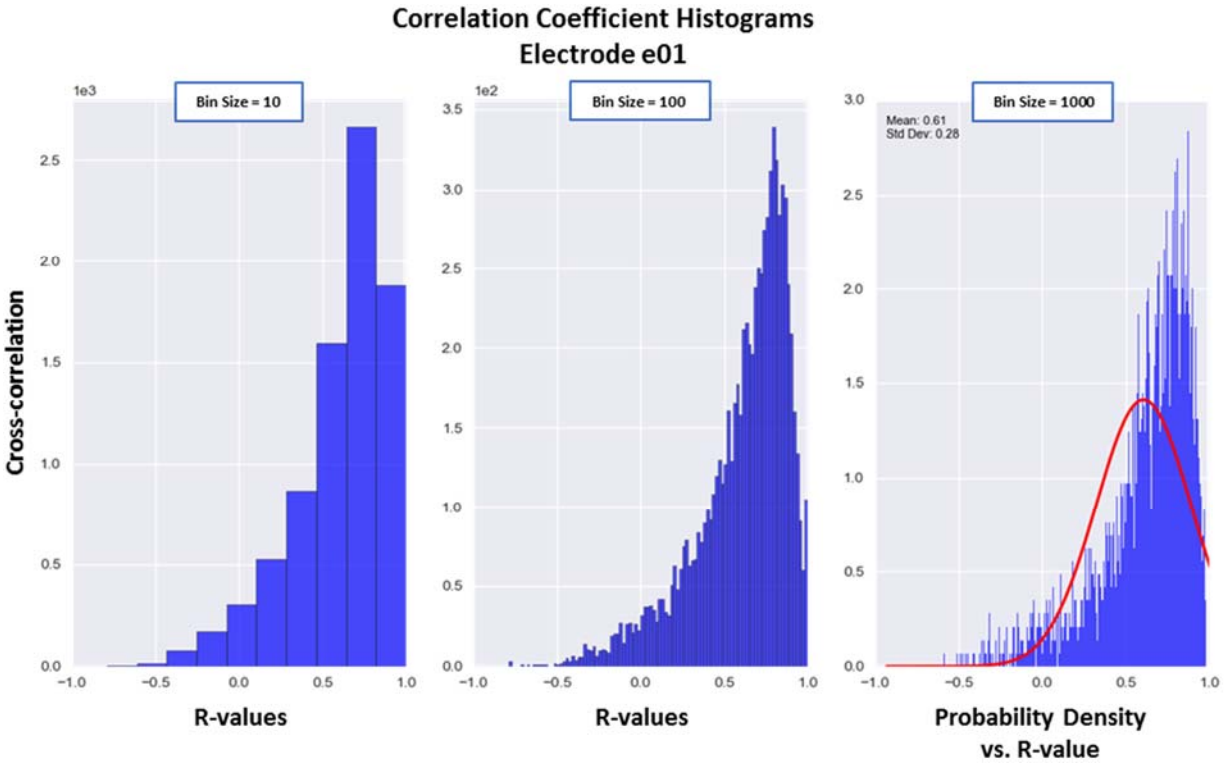

Figure 28 – Histogram plots of various bins sizes, depicting correlaton coefficients for all Utah Array electrodes compared against electrode e01.

In Figure 29 this same KDE is depicted and shown upon each histogram, where all 81 correlation coefficients are plotted for each electrode shown. Throughout this time sequence of 1,000 computed data points, the correlation coefficient results in positive probability densities, indicating that with one electrodes voltage change, so also the other electrodes respond accordingly. The results range from 0.39 – 0.74, where the lower results belong to electrodes e44 at 0.39, e50 at 0.44, and electrodes e32, e58, e60 at 0.51.

Figure 29 – Histogram plots of correlaton coefficients for all Utah Array electrodes compared against each of the specified electrode positions.

LAG PLOTS

A lag peak is a cross-correlation plot feature, that are local maximas of the cross-correlation function at $\tau \neq 0$. More explicitly, a lag peak indicates at what lag time, $\pm\tau$, that the primary electrode signal is similar to the secondary electrode signal. For example, the auto-correlation function of Figure 24 depicts for electrode e01, 2 lag peaks at either side of $\tau = 0$. Figure 30 shows a close-up view of these negative lag peaks at $\tau = -146$ mSec and $\tau = -322$ mSec. Equally, Figure 31 gives another close-up view of such cross-correlation results pertaining to electrodes e01 x e33, just as is shown in Figure 25 where $\tau = -115$ mSec and $\tau = -262$ mSec. The importance of a lag plot in this research study is that of its importance as a building block into the forthcoming, non-traditional, analysis upon the Utah Array electrode signals.

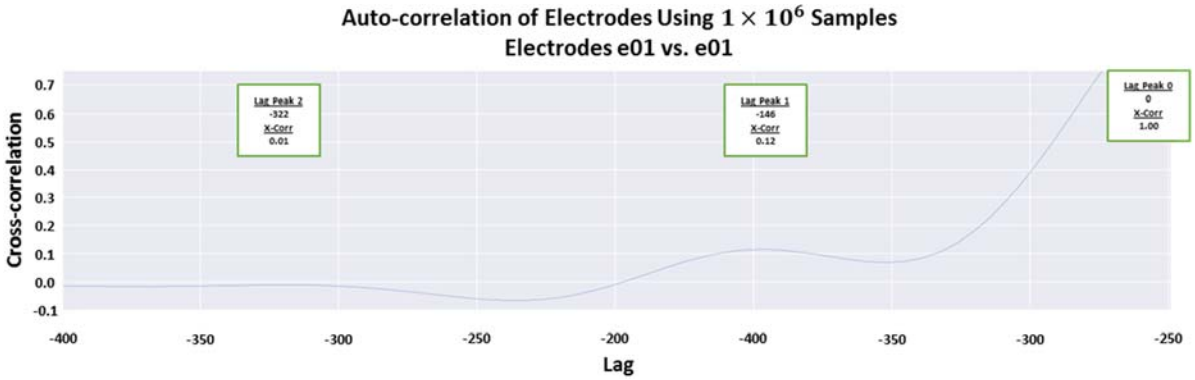

Figure 30 – The auto cross-correlation lag plot of – τ for electrode e01, where $\tau_{peak\_1} = -146$ mSec and $\tau_{peak\_2} = -322$ mSec.

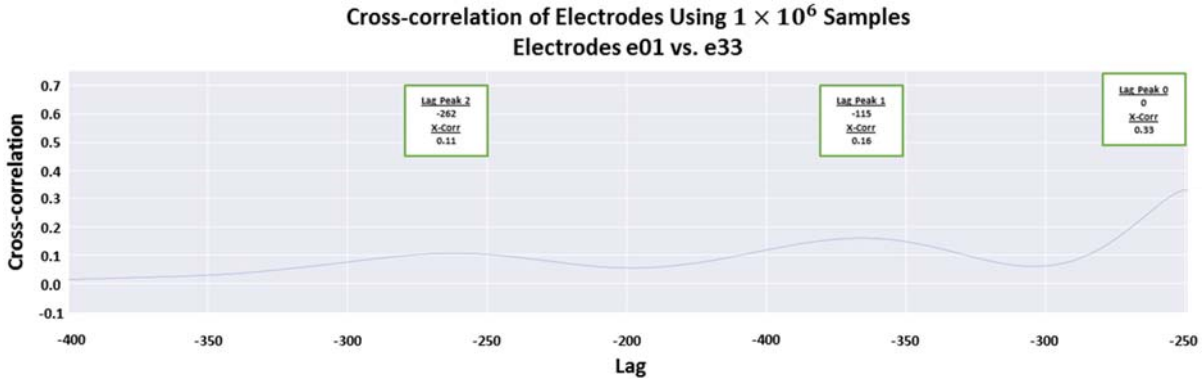

Figure 31 – The cross-correlation lag plot of – τ for electrode e01 x e33, where $\tau_{peak_1} = -115$mSec and $\tau_{peak\_2} = -262$ mSec.

NON-TRADITIONAL ANALYSIS

A signal analysis approach pertaining to tradition is classically expected to include elements of statistics. In this way, the progress thus far reported pertains to only one of tradition. However, for the proceeding sections this investigation is expanded to include non-traditional tools such as Utah Electrode Array time series animations, for example, by using heatmap and histogram plots of electrode voltages and by also plotting their differences and via animated histogram plots. Both correlation coefficients and cross-correlation lag peaks are animated into heatmap plots, where yet another investigation is to establish electrode equivalency classifications via visual interpretation of electrode cross-correlation lag peak differences. Doing so, making equivalency classifications, is a direct consequence of the first novel approach used in this paper (below) for time series observations of pair-wise cross-correlated electrodes, where the electrode lag peak differences are computed and plotted at each time sample. The second novel approach used is to evaluate the same electrode time series plots of cross-correlations that further deduce point-slope magnitudes between each of the lag peaks, and to determine these

differences as well. In the following sections of this paper, each of these methods are discussed and demonstrated. As before, these analytic methods are intended to develop improved research results in search of neural network patterns.

ANIMATIONS

VOLTAGE HEATMAP PLOTS

Animations of interest to this study entail heatmap plots used primarily for visualizing various perspectives of the Utah Electrode Array elements across a specified period of time. The physical geometry of the Utah Array is identical to the heamap plot starting with electrode e01 in the upper-left position, and moving row-wise to the bottom-right position where electrode e81 is located. With this, the most immediately available animation to plot is that of each electrodes voltage at a single time step, then, increment to the next time step in the same way. To demonstrate this, using Figure 32 as time sample 1 and Figure 33 as time sample 2, where this sequential animation time step is established in this manner. Observe that each of the heatmap plots (including the animations) are produced in both false color and in gray scale to offer dual methods of data visual interpretation. Figure 34 gives a visual comparison of the differences between individual electrode voltages with between Figure 32 and Figure 33.

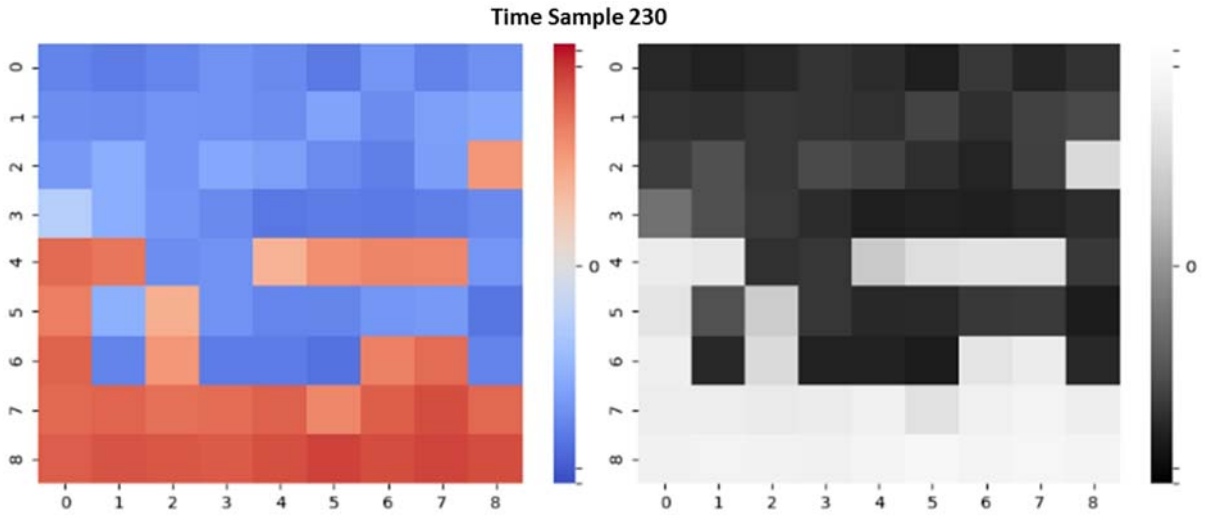

Figure 32 – Heatmap plots depicting a single time step in Utah Array electrode voltages, shown in both false color and gray scale for improved interpretability.

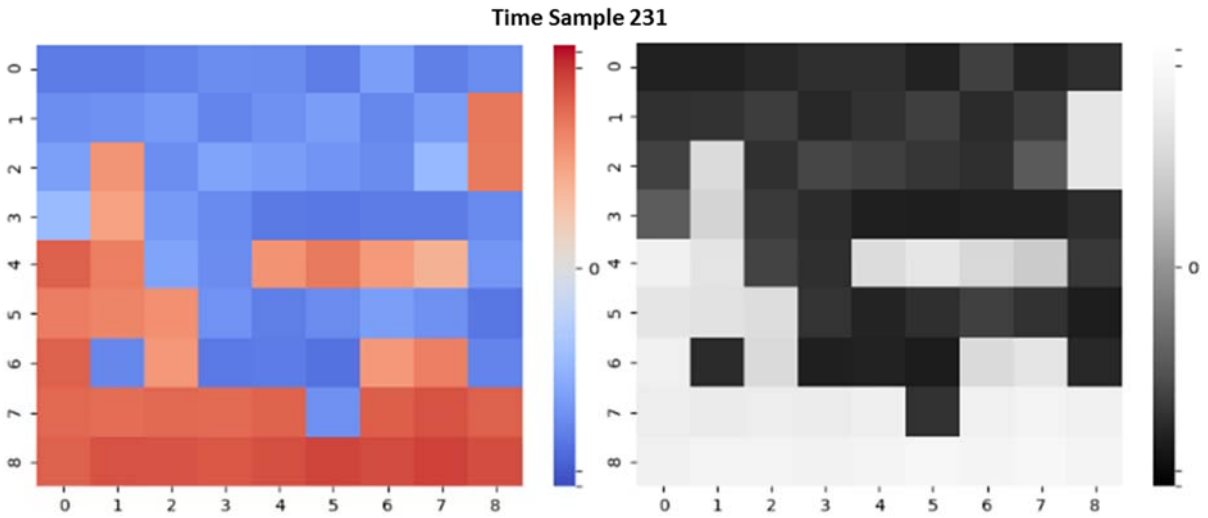

Figure 33 – Heatmap plots depicting the next time sample (after the above figure) of the Utah Array electrode voltages, shown in both false color and gray scale for improved interpretability.

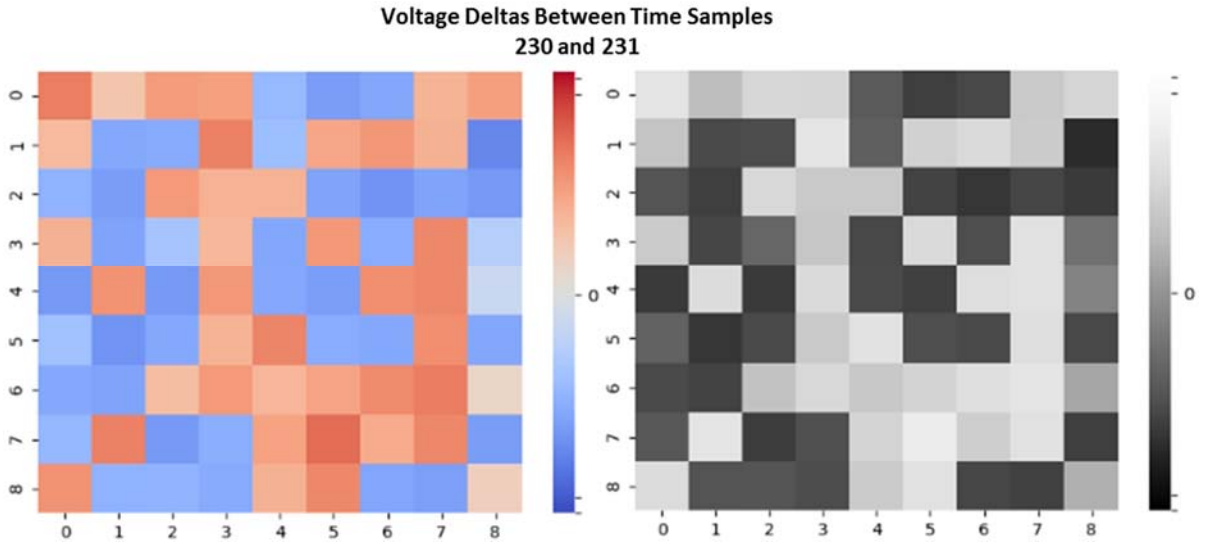

Figure 34 – Heatmap plots depicting the differences in Utah Array voltages from the time step shown in the above 2 figures.

VOLTAGE HISTOGRAMS

The electrodes of the Utah Electrode Array can also be animated via a histogram plot, similar to what is shown in Figure 35, below, where 3 relatively random time points are given. For this figure, each of the time samples shows all 81 electrode voltages at the specified time sample of the histogram plot. An animation then, in this fashion, shows a wide variability in electrode voltages while also returning the additional benefit of a capability to identify outliers. For maximal interpretation of such plotted results, the combination of both a heatmap and a histogram animation offers improved interpretation of the Utah Electrode Array characteristics.

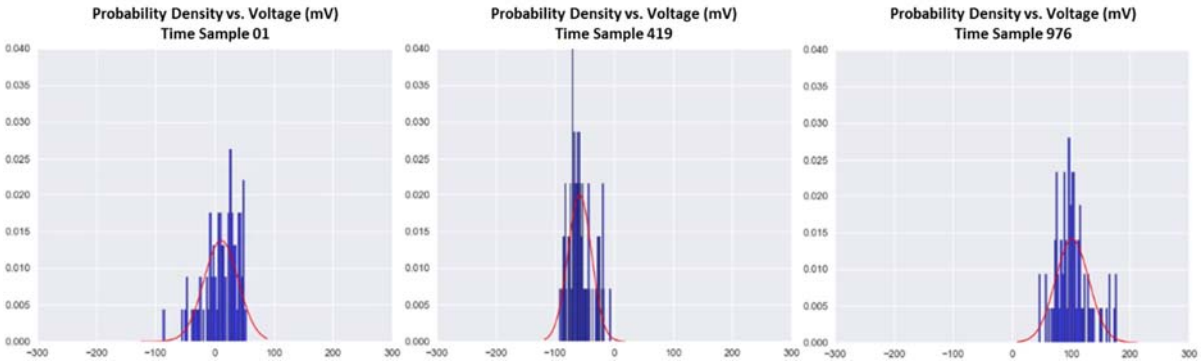

Figure 35 – A selection of histogram plots to demonstrate the extent of varying electrode voltages across time.

CORRELATION COEFFICIENT HEATMP PLOTS

An additional perspective to build upon Figure 32 and Figure 33, where electrode voltages are introduced as animations, is to establish correlation coefficients in like manner. Doing this allows for improved investigation outcomes regarding electrode-to-electrode relatedness. For example, Figure 36 and Figure 37 show a heatmap plot in a 2-frame time-sequence, that several of the plotted coefficient results are negatively correlated. For each plot (both in false color and gray scale) for example, the interpretation of electrode e01 as the left-most column correlated with electrodes e28-e34 results in the negative coefficients shown in blue (or dark gray shading for the gray scale plots), while most other electrodes increase in their coefficient values for the given time sequence. The average change in state between time steps of the Utah Array, trending from lower to higher coefficient magnitudes, is an interesting state of dynamic change as events such as this possess the potential to help unlock any unseen neural network activities from within the Utah Electrode Array dataset. Note the plots below use a log function as a Look Up Table transformation, for improved visual contrast.

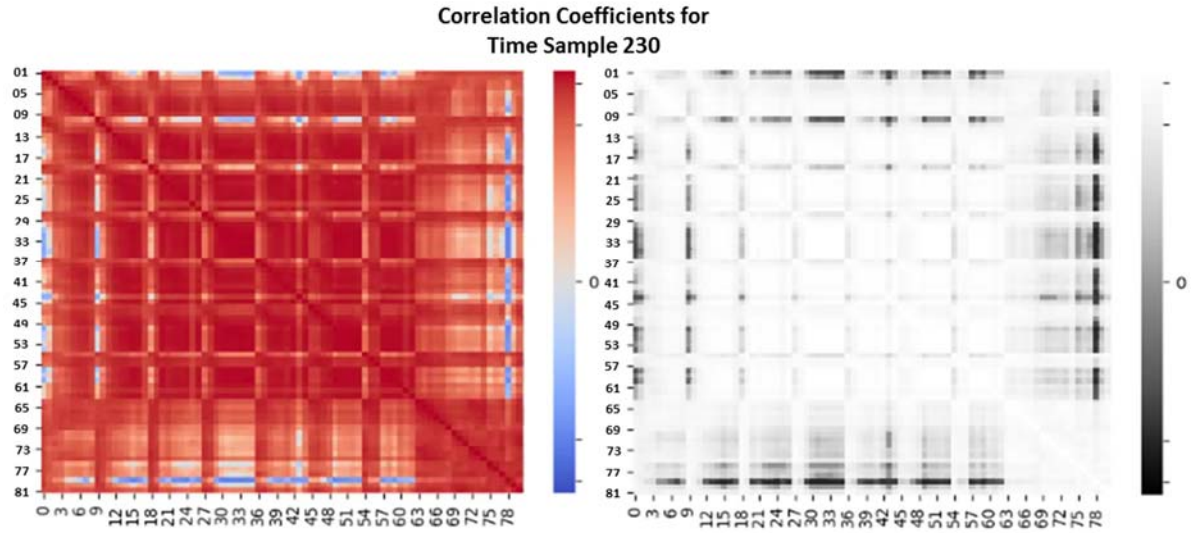

Figure 36 – Heatmap plots in both false color (left) and gray scale (right) showing the first of 2 sequential correlation coefficient results of a specific time sample.

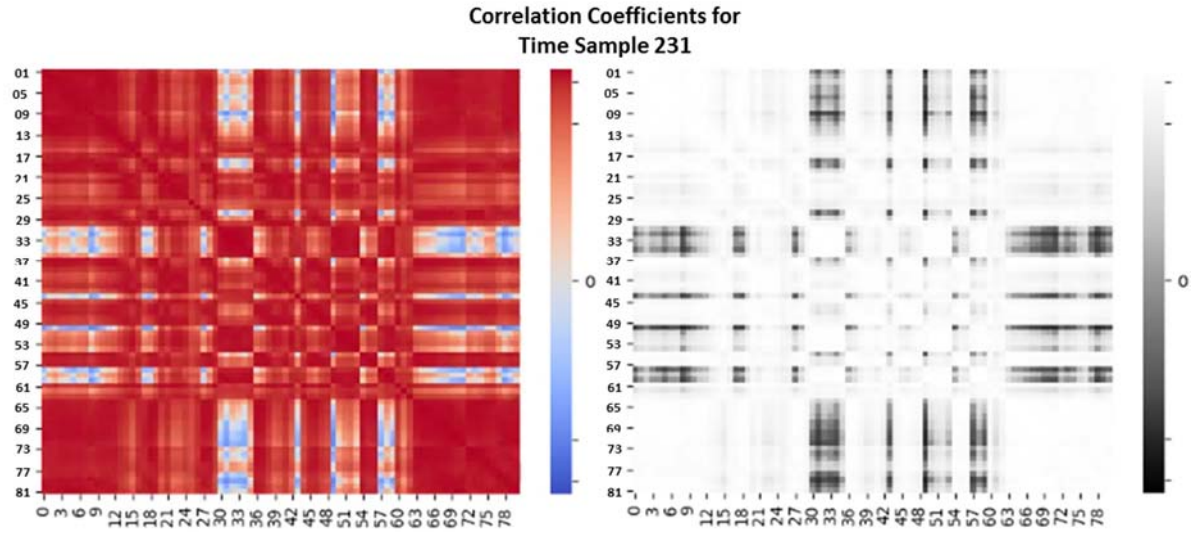

Figure 37 – Heatmap plots in both false color (left) and gray scale (right) showing the second of 2 sequential correlation coefficient results of a specific time sample.

LAG PLOTS

As previously discussed for a lag plot, the cross-correlation function is used to establish its associated lag peaks, or local maximas, located at either side of $\tau \neq 0$. The same constructs are prescribed here for the intention of similarly animating a lag plot time sequence. In this case, Figure 38 and Figure 39 both show heatmap plot results specific to lag peak animations, where the time of the chosen peaks are aligned with those shown in Figure 30 and Figure 31, where $\tau = (-115, -262)$ mSec. To illustrate the underlying methodology used for this animation, 3 peaks are plotted in series: 'Lag Peak 0' indicates the

cross-correlation magnitude at $\tau = 0$, 'Lag Peak 1' is that of $\tau = -115$ mSec, and 'Lag Peak 2' gets $\tau = -260$ mSec. The latter 2 lag times are calculated and chosen as an averaged value of all 81 electrode cross-correlation lag peaks, for both specified peak 1 and peak 2 lags, per electrode e01. As such, for the two figures below, each of the plots gives cross-correlation magnitudes for each of the averaged lag peaks, τ. To compare Figure 36 and Figure 37 with Figure 38 and Figure 39, the 2 pair of figures are in agreement in that the plotted cross-correlation lag peaks reflect a similar increase in correlation coefficient values. This is observed between the 'Lag Peak 1' and 'Lag Peak 2' plots of Figure 38 and Figure 39, particularly for the first and last 3 rows, and equally so upon the first column of both Figure 36 and Figure 37.

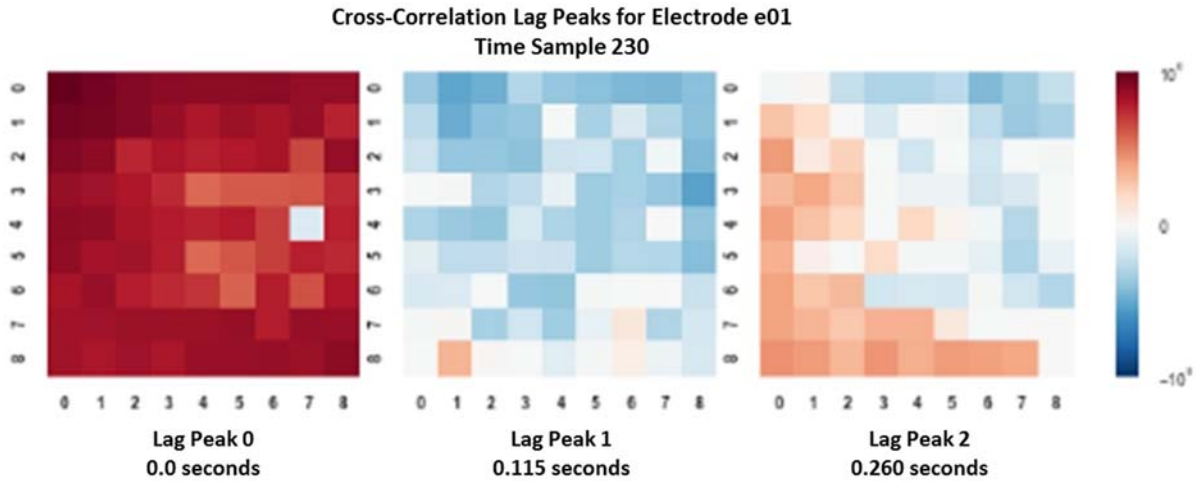

Figure 38 – Heatmap plots of specific correlation lag times (peaks) showing the first of 2 sequential cross-correlation results.

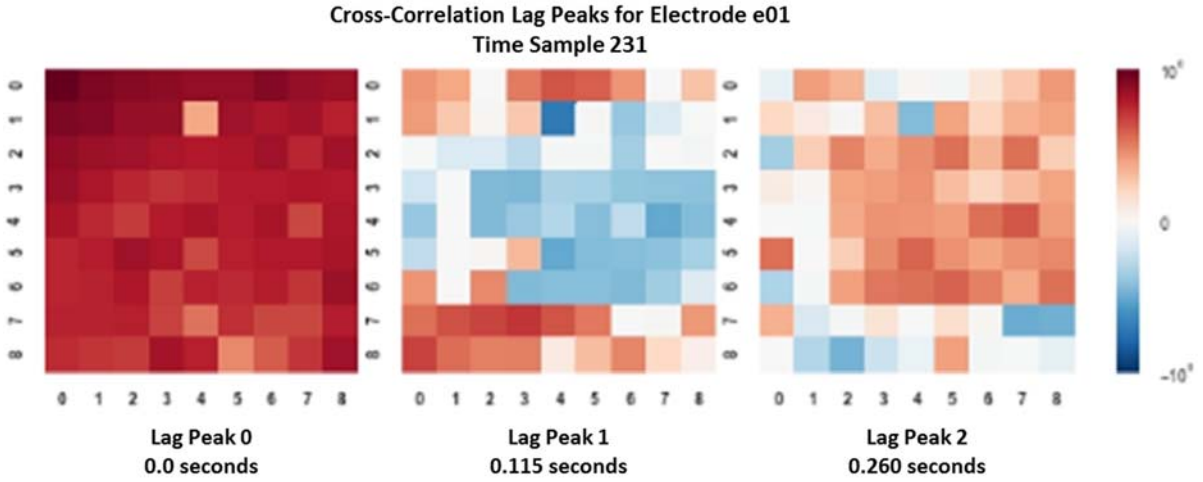

Figure 39 – Heatmap plots of specific correlation lag times (peaks) showing the second of 2 sequential cross-correlation results.

TIME SERIES: CROSS-CORRELATION DIFFERENCES

A close-up investigation was conducted to observe each of the 3 lag peak relationships between the cross-correlation values at $\tau_0 = 0$ mSec, $\tau_1 = 115$ mSec, and $\tau_2 = 260$ mSec pertaining to the 1-million data points of the Utah Array dataset. The next several figures, below, show the results of the cross-correlation function as repetitively calculated using a bin size of 1,000 data points between the 2 specified electrodes. For each of these resultant functions, the 3 lag peak magnitudes are plotted to their associated time series charts in pair-wise combinations. The easiest visual interpretation of this is shown in Figure 40, where the auto-correlation function results of electrode e01 are given. The upper plot gives both primary electrode e01 as a blue trace at $y = 1.0$ and secondary electrode (also e01) as a green trace. The y-axis label indicates that lag peaks τ_0 and τ_1 are plotted. Identically, the remaining 2 time series plots are shown for the pair-wise plotting of (τ_0, τ_2), and (τ_1, τ_2) from top to bottom.

One additional feature of the following 3 figures, as shown in the third row of Figure 40, are red triangle markers positioned at the bottom of the time series plots shown. These markers indicate which of the data points from the plotted cross-correlation results meet the threshold as shown within the plot title. For Figure 40, this threshold is 0.023, and is concluded as the difference between lag peak values at each plotted lag, τ. This is to say, that for each instance that the specified pair-wise lag peak difference is less than this threshold, Δ_{Xcorr}, a triangle marker is placed onto the plot. Observe that these markers are tabulated to the right of the plot, showing exact lag times and the cross-correlation results used for computing the cross-correlation difference. Adding complexity to this framework, Figure 41 demonstrates for electrodes e01 x e33 the same implementation showing an increased magnitude of triangle markers as compared to the prior. Increasing this threshold consequently increases the number of triangle markers plotted. This is exemplified in Figure 42 where both the magnitude of resultant triangle markers and tabulated results, for the same e01 x e33 cross-correlation result, reflects an increase in lag peak values meeting the criteria.

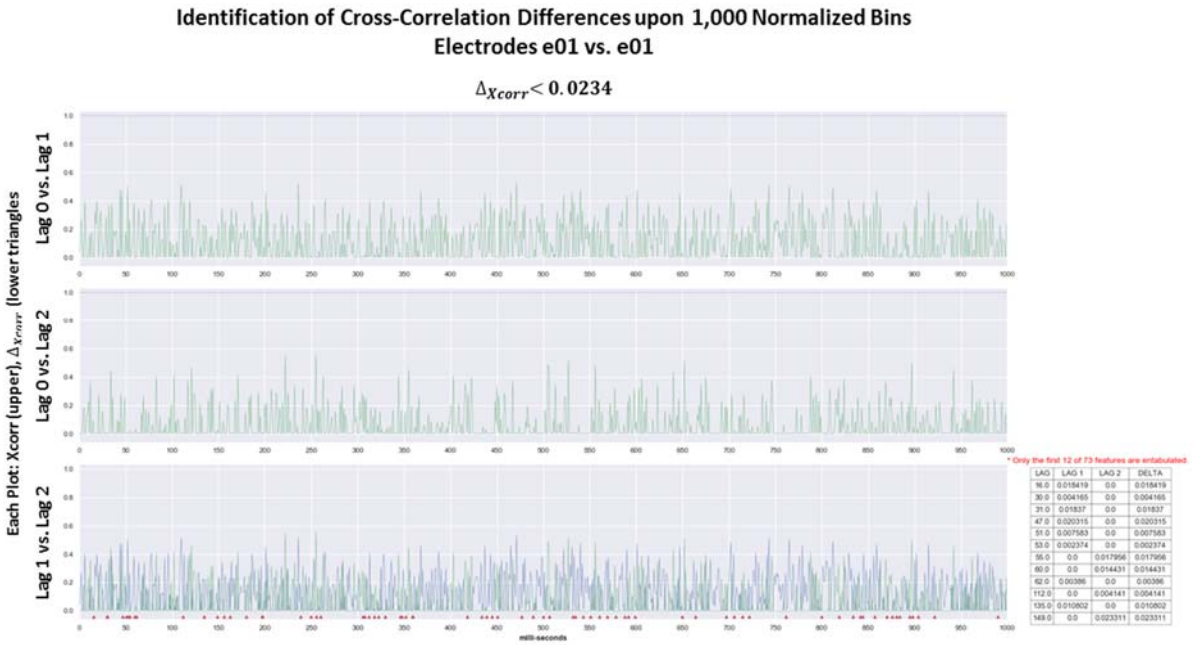

Figure 40 – Auto-correlation results for electrode e01 x e01, where the qualifying differences between 3 pair-wise lag peaks are concluded with red marker triangles and associated tabulation. Note: threshold < 0.0234.

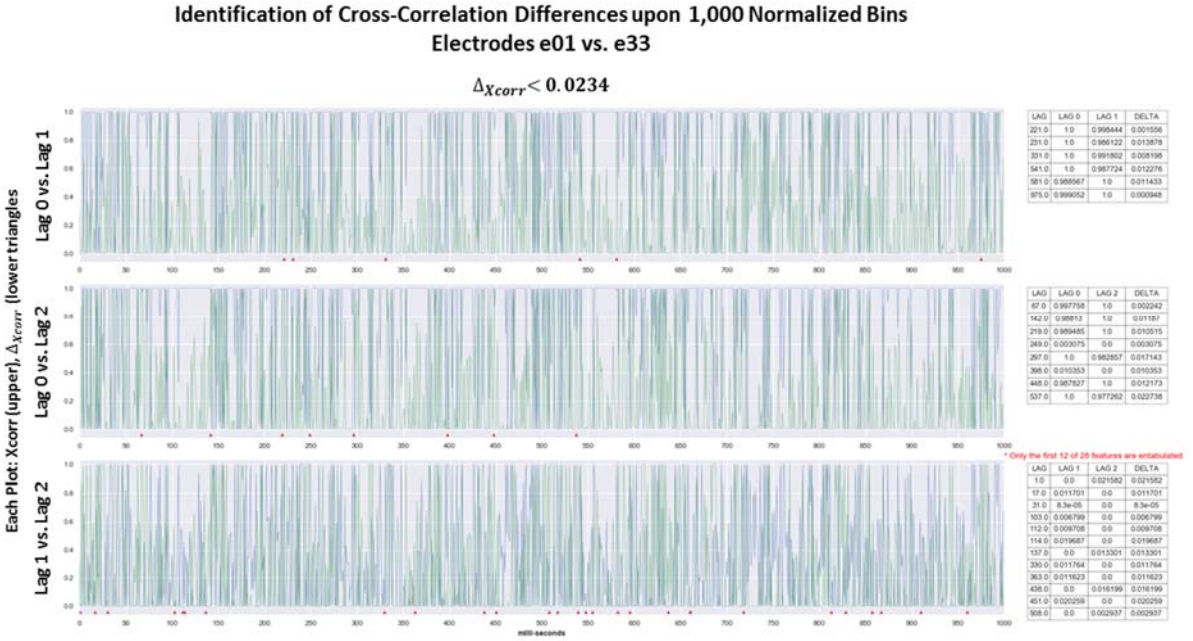

Figure 41 – Cross-correlation results for electrode e01 x e33, where the qualifying differences between 3 pair-wise lag peaks are concluded with red marker triangles and associated tabulation. Note: threshold < 0.0234.

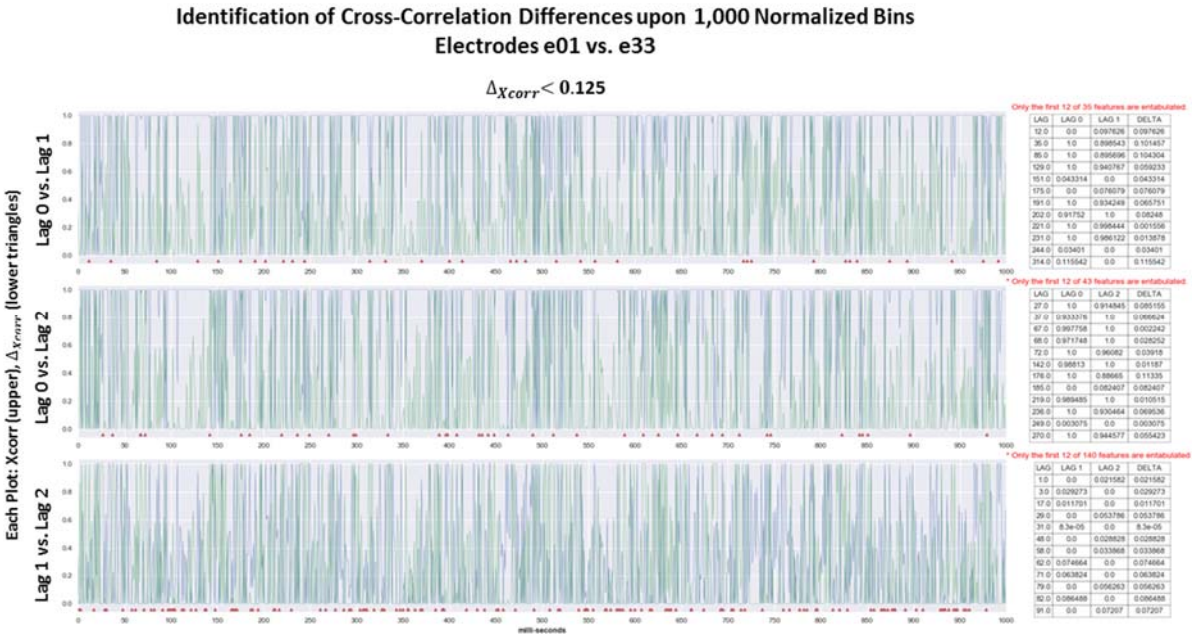

Figure 42 – Cross-correlation results for electrode e01 x e33, where the qualifying differences between 3 pair-wise lag peaks are concluded with red marker triangles and associated tabulation. Note: threshold < 0.125.

The principle of visualizing these differences is further abstracted as shown in

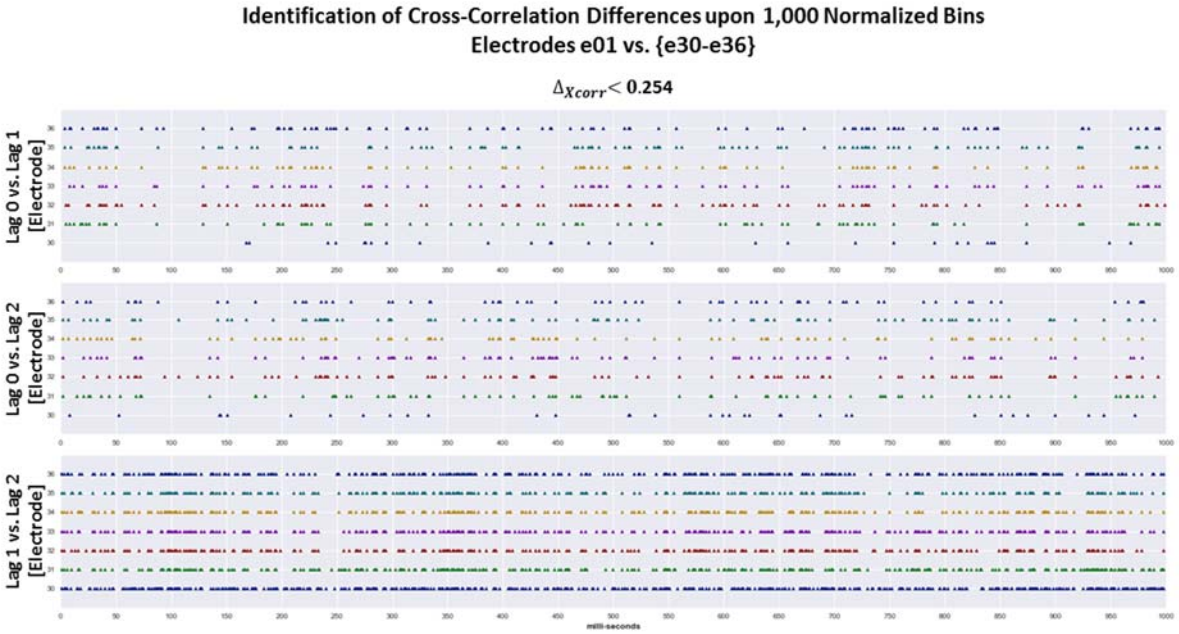

Figure 43 – A time series plot of a smaller subset of electrodes vs. time, showing the result from cross-correlation differences in comparison to the 3 lag peaks at previously specified times.

below. In this new time series plot, where $\Delta_{Xcorr} < 0.254$, only the aforementioned triangle markers, the cross-correlation differences, are plotted. Here, only a subset of electrodes are

shown upon the y-axis, giving light to this investigations purpose to similarly group related electrodes into equivalency classifications. The expectation for employing this type of plot then is for improved pattern recognition of both the underlying neuronal activity and its associated dataset. In

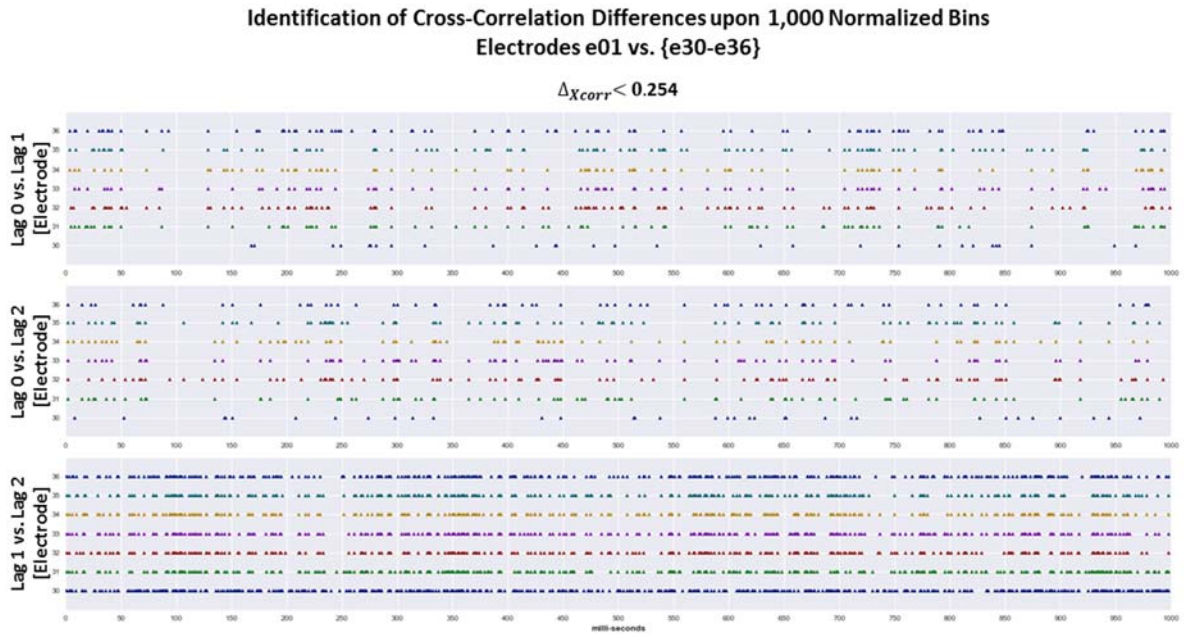

Figure 43 – A time series plot of a smaller subset of electrodes vs. time, showing the result from cross-correlation differences in comparison to the 3 lag peaks at previously specified times.

showing 7 electrodes, the 3rd time series plot exemplifies this very perspective. However, such a pattern may not be as clearly shown as it is in the 52-electrode results given in Figure 44. Comparing this figure to the prior indicates a larger subset of Utah Array electrodes used to show a better defined pattern-like result. From this new perspective, one might expect to observe neuronal patterns not only by the vertical alignment of markers at a given time sample, but also by a horizontal sequence of markers that contribute to pattern-like activity identification.

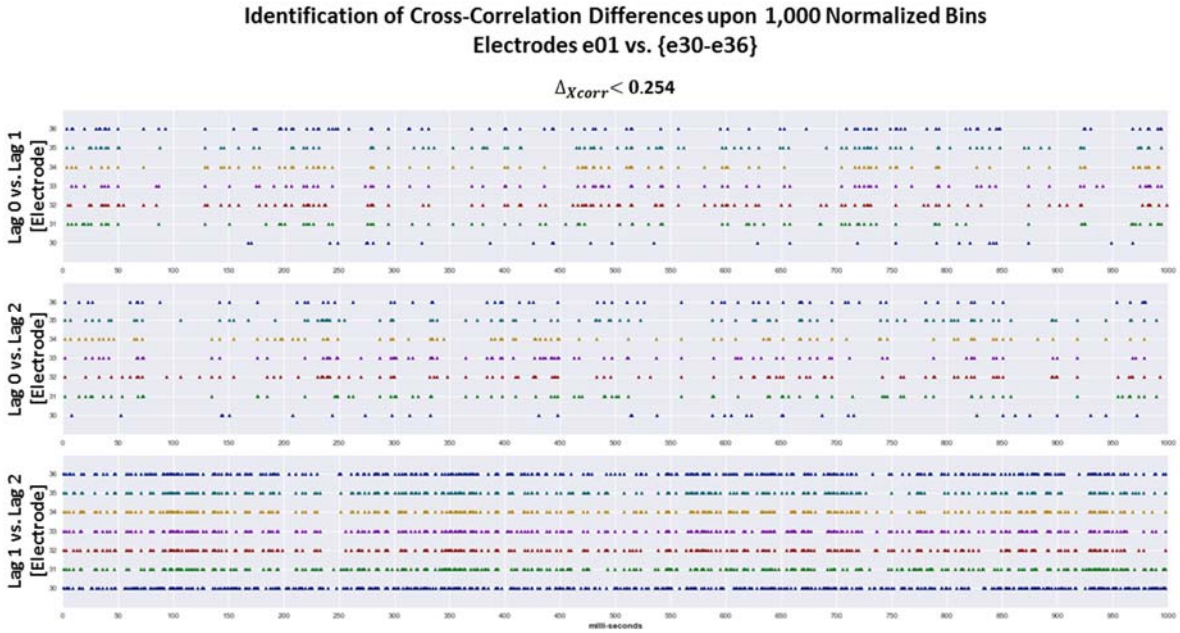

Figure 43 – A time series plot of a smaller subset of electrodes vs. time, showing the result from cross-correlation differences in comparison to the 3 lag peaks at previously specified times.

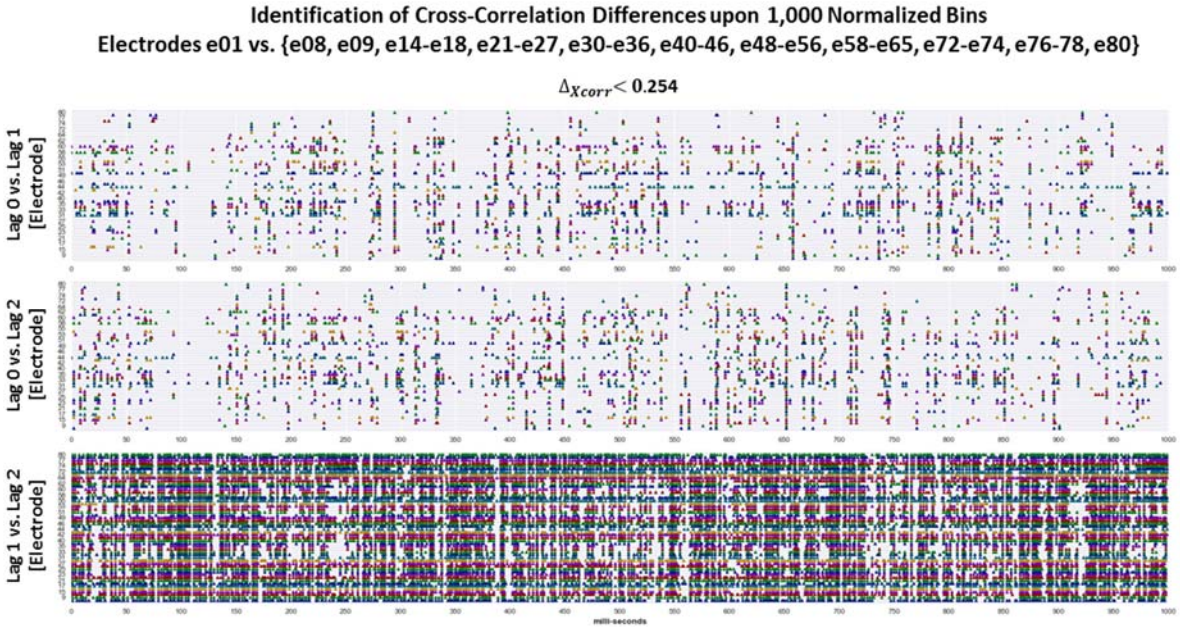

Figure 44 – A time series plot of a larger subset of electrodes vs. time, showing the result from cross-correlation differences in comparison to the 3 lag peaks at previously specified times.

TIME SERIES: CROSS-CORRELATION POINT-SLOPES AND DIFFERENCES

So as to further investigate the pattern-like characteristics of the Utah Electrode Array dataset, a new type of time series plot is established to extend the principles of cross-correlating and said lag peak differences as previously discussed, to visualize slopes of the point-to-point cross-correlation lag peak results. In this way, each of the time series plots

shown below in Figure 45 pertain to a single lag peak, τ, such that $\tau_0 = 0$ mSec, $\tau_1 = 115$ mSec, and $\tau_2 = 260$ mSec (from top to bottom). The slope calculation between time samples is developed in the standard way, mainly:

$$slope = \frac{\Delta y}{\Delta x}. \tag{14}$$

Such point-slope results appear in the $y = 0$ axis of these time series plots below (Figure 45, Figure 46, and Figure 47). The thick black trace in these plots gives the primary electrode of the cross-correlated time samples. Below this central axis then, is an additional plot

feature where triangle markers are placed, to depict the relative difference between the sequential point-slope magnitudes.

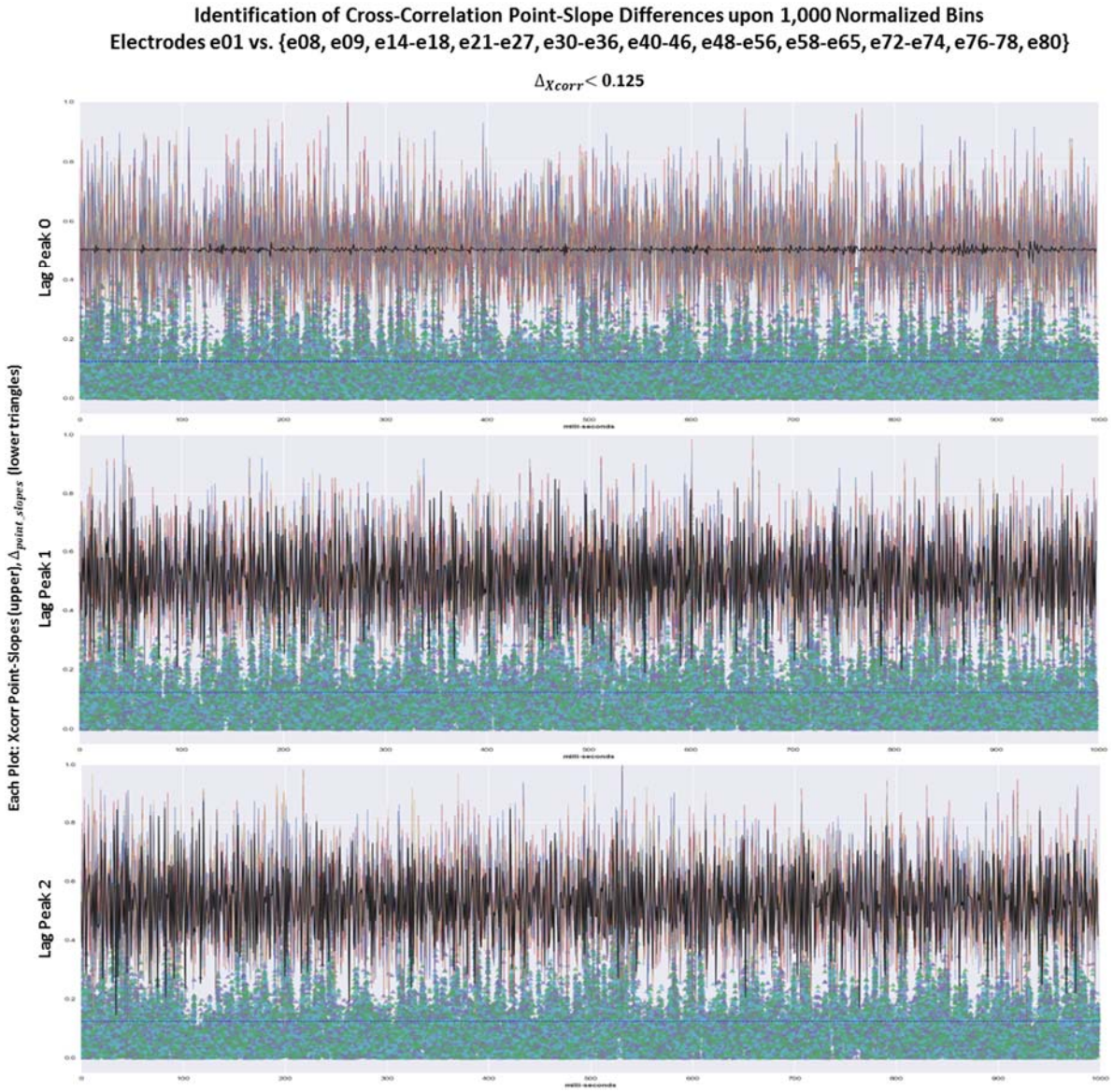

Figure 45 – Point-slope and point-slope differences plots for numerous electrode cross-correlation magnitudes.

It is advantageous to demonstrate this new plot type on a shorter time scale, just as Figure 46 shows, and by also plotting with fewer electrodes, as shown in Figure 47. A shorter time scale such as this gives improved delineation between time steps over the prior figure, and by implementing fewer point-slope traces gives a more useable resolution of data points. Here, electrode e01 is cross-correlated with the others, and when the point-slope differences are consecutively small (say, sequentially less than the plotted horizontal blue line), it is

plausible that a recurring pattern might emerge as observable in this way. Indeed, such a pattern is shown for time samples 215 thru 220, or 226 thru 228, particularly visible in Figure 47.

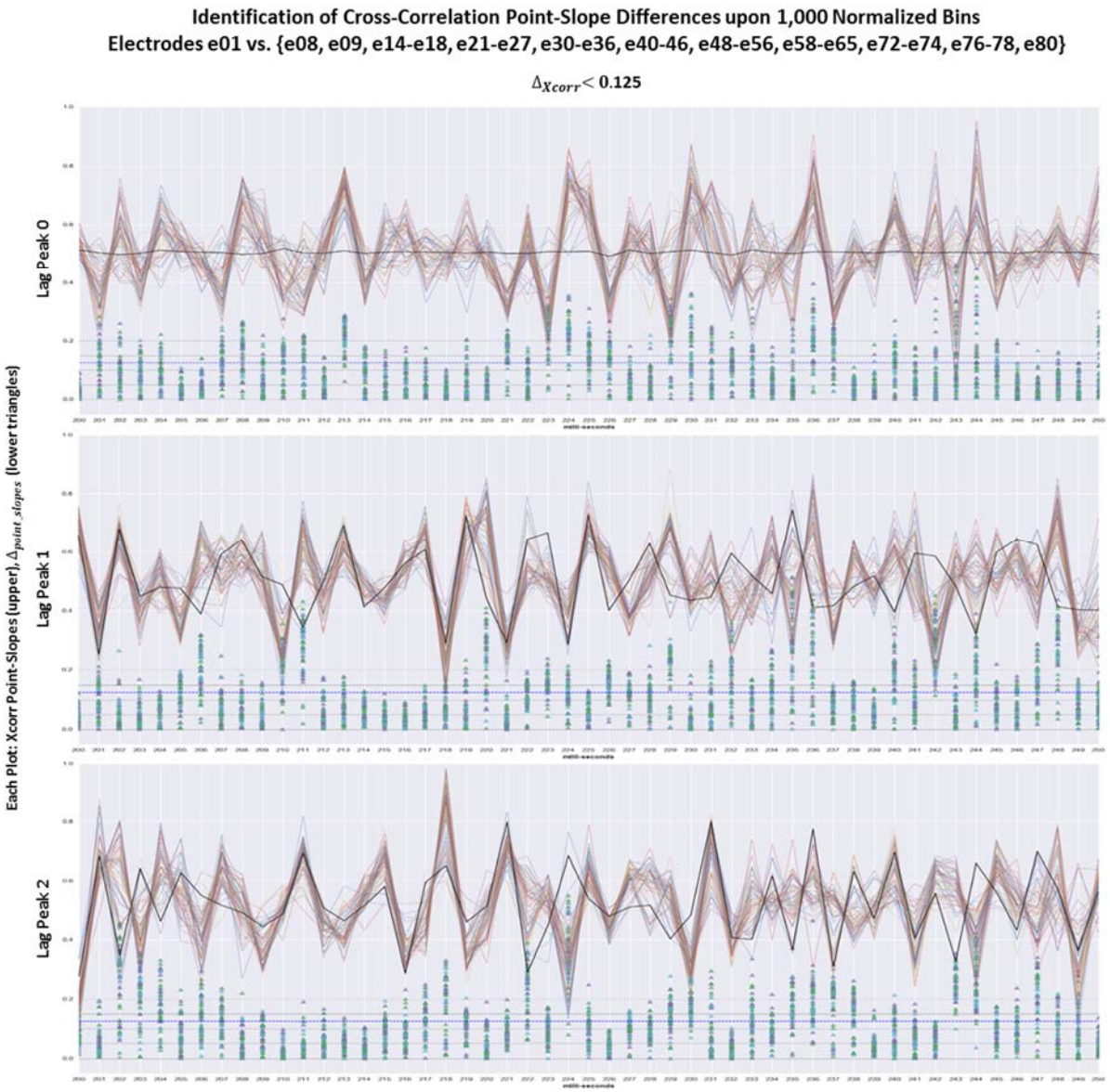

Figure 46 – Point-slope and point-slope differences plots for numerous electrode cross-correlation magnitudes, for a restricted time scale of 200 – 250 mSec.

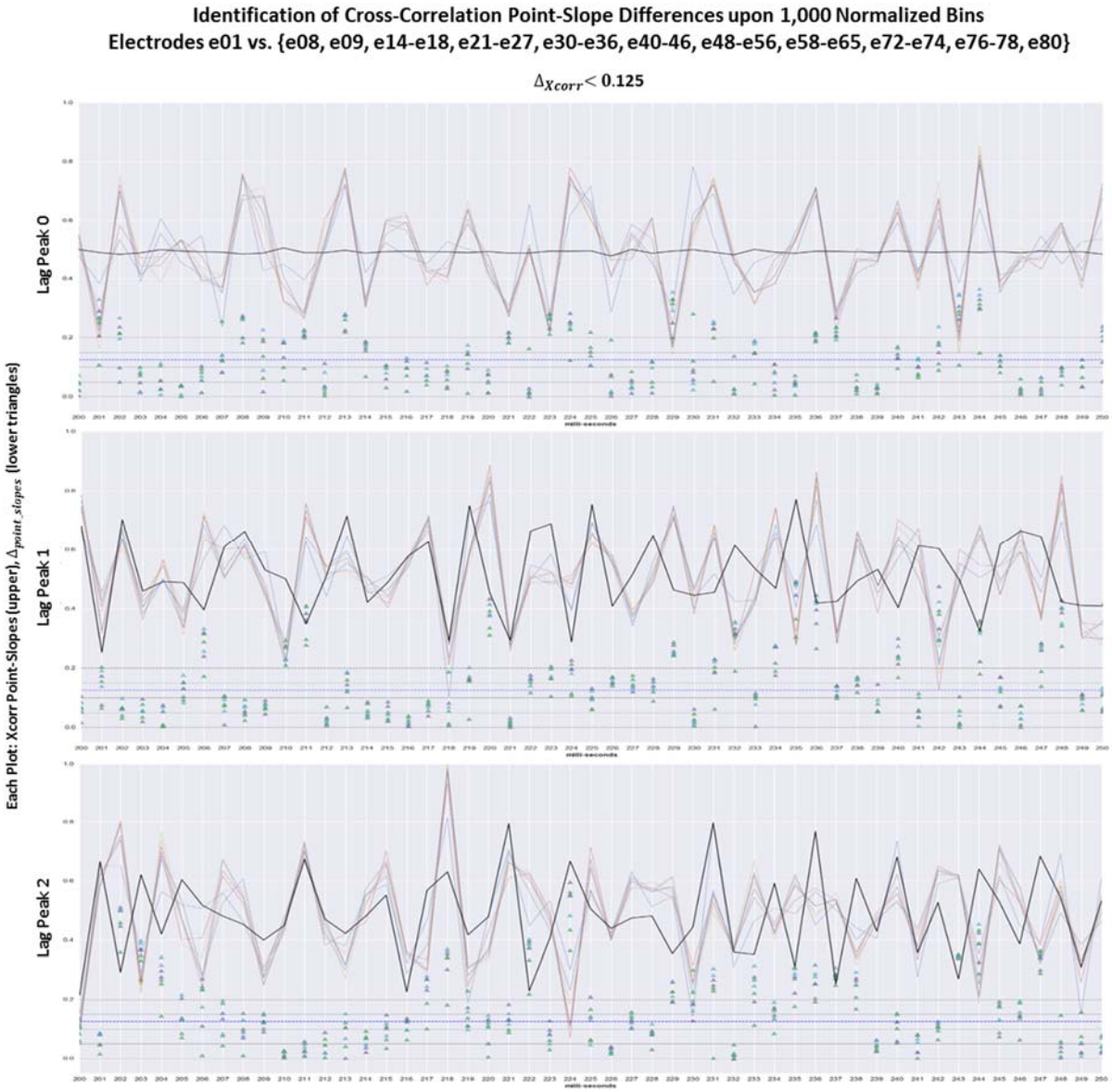

Figure 47 – Point-slope and point-slope differences plots for fewer electrode cross-correlation magnitudes, for a restricted time scale of 200 – 250 mSec.

RESULTS

One of the most basic tools for employing statistical analysis upon the Utah Electrode Array in this paper is the scatter plot, as it is a critical factor in substantiating correlation results between electrode constituents. As demonstrated in Figure 20 and Equation (8), such correlations are interpreted as both pearson-R values and correlation coefficients. In Figure 48 below, the covariance matrix for electrodes e28 thru e36 gives a step-wise progression away from Figure 20 in that scatter plots are clearly embedded within this figure of correlations. Similar to Figure 21 where the basic scatter plot is improved by using density shading for easier

contrast, Figure 48 too shows distinct density characteristics between x-axis electrodes e30 and e31, giving one region that is more highly correlated (rows > 3, electrodes $\geq e31$) and one that is less so (rows > 3, electrodes $< e31$). Specifically, the correlations within each of the individual scatter plots for electrodes e31 thru e36 appear ≈ 1, both by way of the data point distributions and the slope of the regression line, shown in red. This visual demarcation between these contrasting regions of the Utah Array provides direct visibility into those electrodes that share commonalities across time.

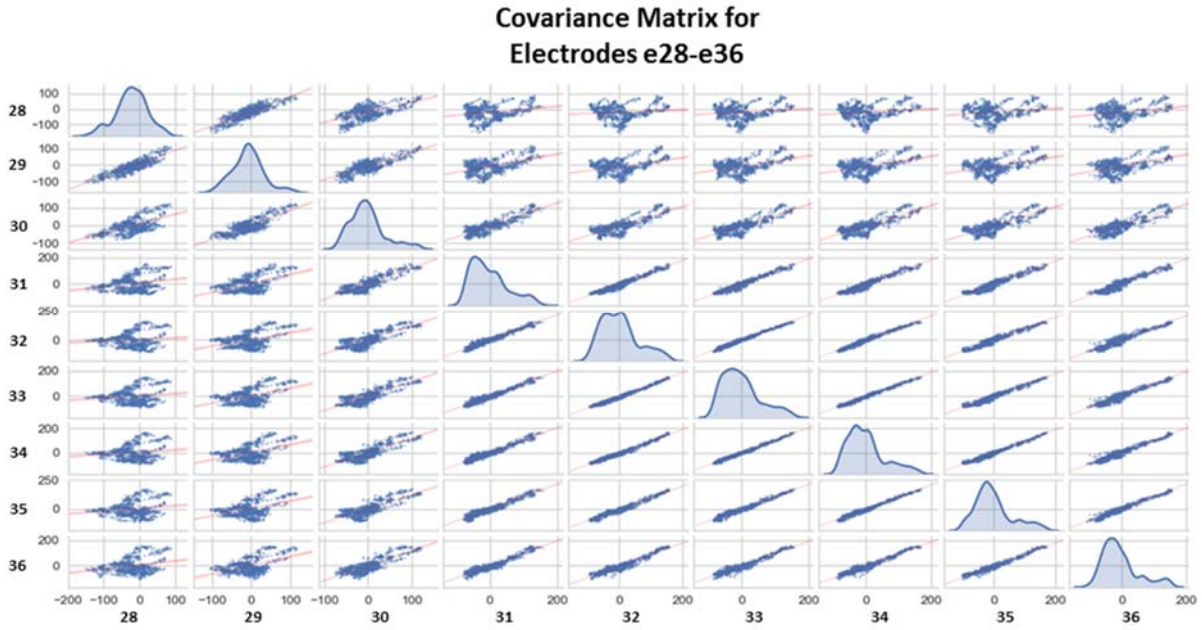

Figure 48 - Correlation Matrix of raw signals, electrodes 28 – 36.

Figure 48, above, is positioned as the 4th red boxed region within Figure 49 (REF: Figure 23) along the diagonal, where a lighter (white) background is seen in comparison with most other plots shown. However, because the latter figure is simply a larger covariance matrix, a scaled-up version using all 81 Utah Array electrodes, patterns similar to those just discussed are visible throughout this larger result. Such patterns appear as rectangular areas of alternating light and dark due to the variance in scatter plot density distributions from plot to plot. To this end, regions of those more highly correlated results are identified in Figure 50, where visible construction lines reflect the method used for the determination of the related electrode subsets, as shown in Table 1 - Results of the covariance matrix giving the most significantly related subsets of electrodes.Table 1 below.

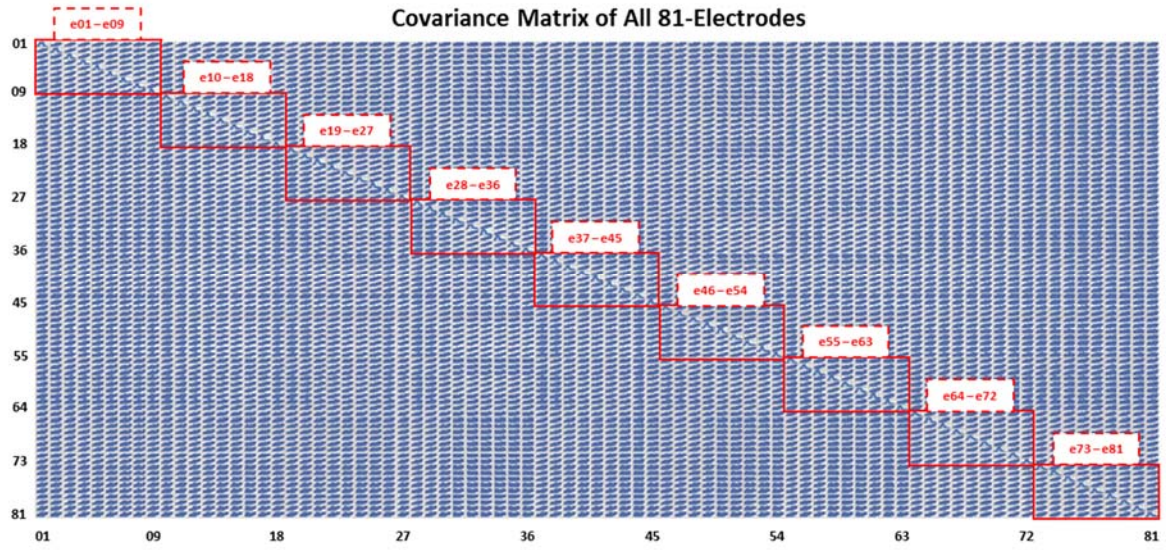

Figure 49 – All outlined sub-sets of correlation scatterplots indicating those evaluated in like manner to Figure 22 and Figure 48 above.

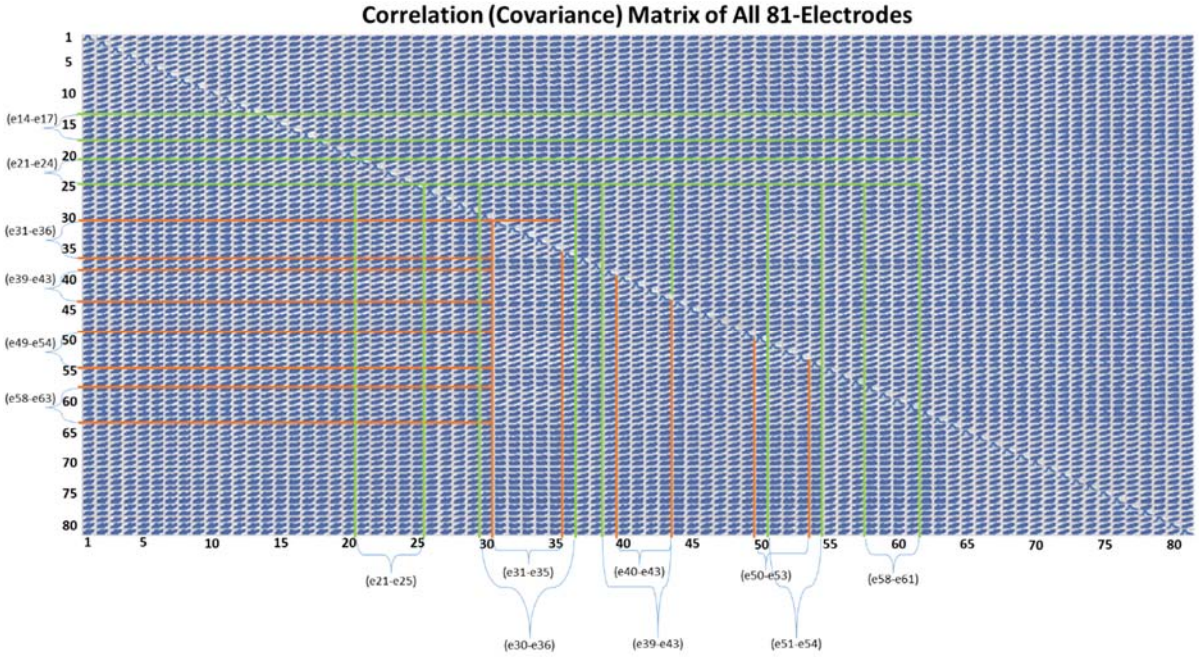

Figure 50 – The covariance matrix plot with the construction lines used to determine those electrodes with inter-dependent relationship.

Table 1 - Results of the covariance matrix giving the most significantly related subsets of electrodes.

y-axis Electrodes	x-axis Electrodes Covariance				
e14-e17	e21-e25	e30-e35	e39-e43	e51-e54	e58-e61
e21-e24	e21-e25	e30-e35	e39-e43	e51-e54	e58-e61
e31-e36	-	e31-e36	-	-	-
e39-e43	-	e31-e36	e40-e43	-	-
e49-e54	-	e31-e36	e40-e43	-	-
e58-e63	-	e31-e36	e40-e43	e50-e53	-

Another outcome toward the identification of related subsets of electrodes is to use the novel plotting techniques constructed in the last 2 sections - Cross-Correlation Differences, and Cross-Correlation Point-Slope Differences. As previously shown in Figure 40 thru Figure 42, the tabulated cross-correlation results are tied to a threshold variable, Δ_{Xcorr}, that behaves as a filter upon the differences between specified lag peaks, τ, such that only values less than the threshold are returned. In this way, by independently setting the 2 threshold values 0.125 and 0.254, a cross-correlation lag peak comparison of electrode e01 is established across all 81 electrodes of the Utah Array. Such results are shown in Figure 51, where differences are derived for both 1,000 and 2,000 cross-correlation bin sizes. From the figure, these electrode groupings are color coded as GREEN for the smaller bin size, and RED for the larger. The final classification of these electrodes are shown in Figure 52, where Figure 53 confirms this conclusion via a random selection of animated electrode voltage heatmap plots.

Cross-Correlation Differences Results
Electrode e01 vs. All

$\Delta_{Xcorr} = (\tau_0\text{-}\tau_1) < 0.125$

Electrode	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
# of Results	0	0	0	0	0	0	1	4	6	0	0	0	0	11	11	27	9	7	0	1	9	6	26	17	11	18	9	2	1	9	97	125	101	111	112	79	0	2	6	12	8

Electrode	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81
# of Results	6	21	169	19	3	3	6	6	132	63	41	38	67	1	2	4	101	73	100	7	45	11	4	2	3	2	2	0	2	1	0	1	6	3	3	3	4	0	2	2

$\Delta_{Xcorr} = (\tau_0\text{-}\tau_1) < 0.254$

Electrode	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
# of Results	0	0	0	0	1	2	8	12	36	0	0	2	2	42	42	101	27	45	0	1	39	15	82	74	51	97	40	5	5	33	192	250	203	219	232	176	4	7	18	56	40

Electrode	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81
# of Results	21	80	325	75	11	12	27	34	251	152	127	115	158	6	16	15	206	154	216	41	170	46	11	11	10	8	9	3	6	11	11	12	12	11	8	11	17	3	6	6

GREEN FONT
1,000-Bin results

RED FONT
2,000-Bin results

Figure 51 – Results of calculating the differences in electrode cross-correlation lag peaks so as to classify them into specific groups of related voltage characteristics.

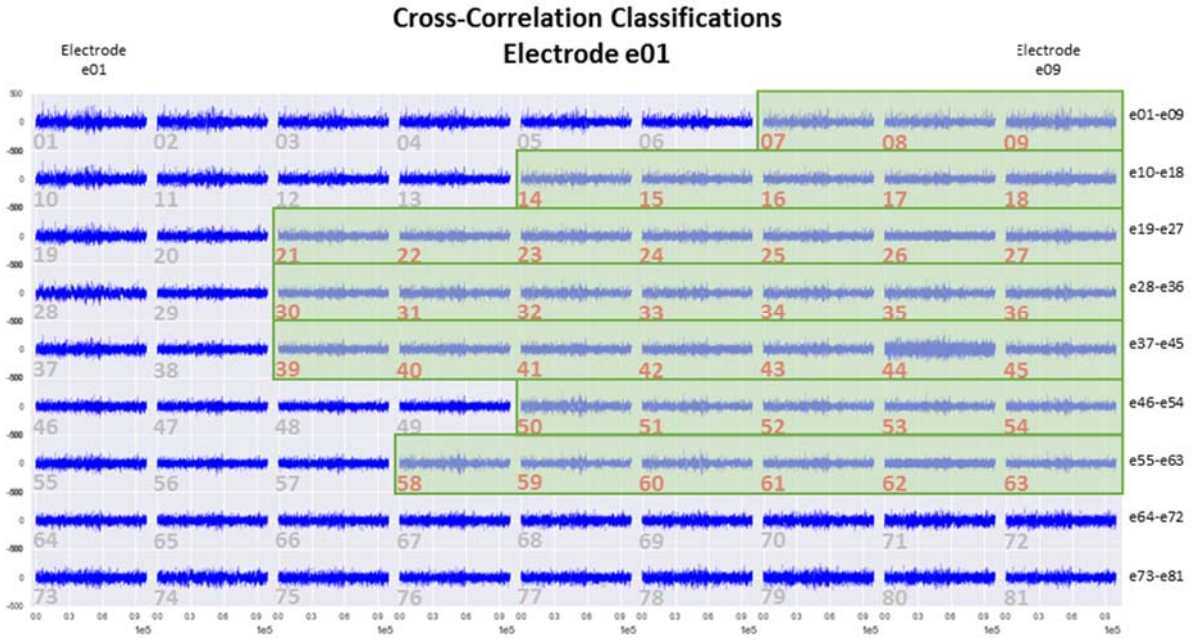

Figure 52 – The geometric arrangmenet of a Utah Electrode Array overlayed with the electrode classifications as derived in Figure 51.

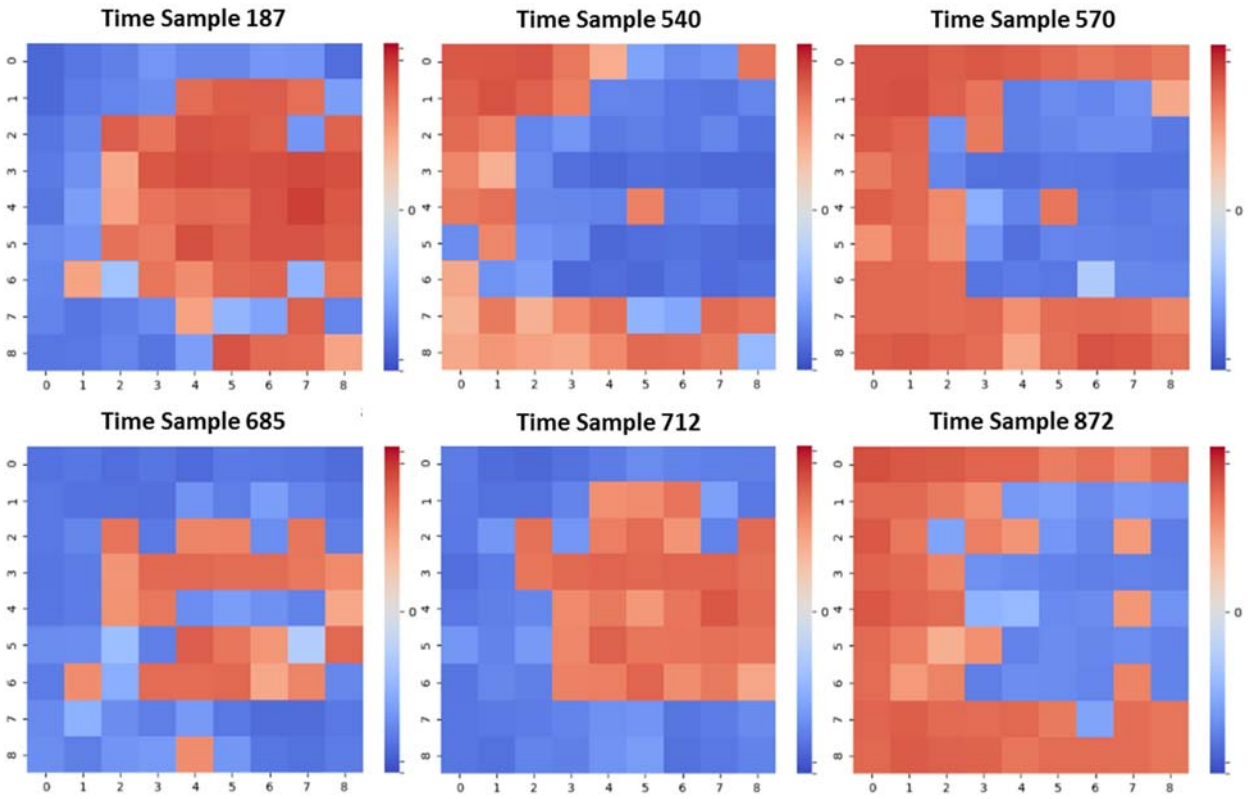

Figure 53 – Electrode voltage heatmap plots of several time samples showing confirmation of Figure 52, where electrode classifications are concluded.

For the cross-correlation point-slope differences plots, as shown in Figure 47 and Figure 54, an additional attempt was intended to further classify electrode characteristics more accurately. In these figures, where an individual triangle is used to indicate the point-slope difference between electrode e01 (shown in black) and the others, the persistence of an electrodes signal is evaluated through various sized windows of time. For example in Figure 54, 2 reference windows are shown between 215-220 mSec and 226-228 mSec where the point-slopes are relatively similar, as shown by the clustered results at each of these time steps. Foundationally, each of the clusters provides a relative ranking of cross-correlation lag peak values with respect to electrode e01, that from moment-to-moment have similar slope. Therefore, the time sequenced electrode point-slopes and their associated differences suggest that the smallest resultant difference is the specific electrode that most closely shares its signal with electrode e01, from one time step to the next. Hence for a given window of the sequential clustering of point-slope differences beneath a threshold, such as < 0.125 (blue dashed horizontal line,

below), a more accurate observance of the nearest-valued electrode is concluded, reflecting the dominant electrode with respect to e01.

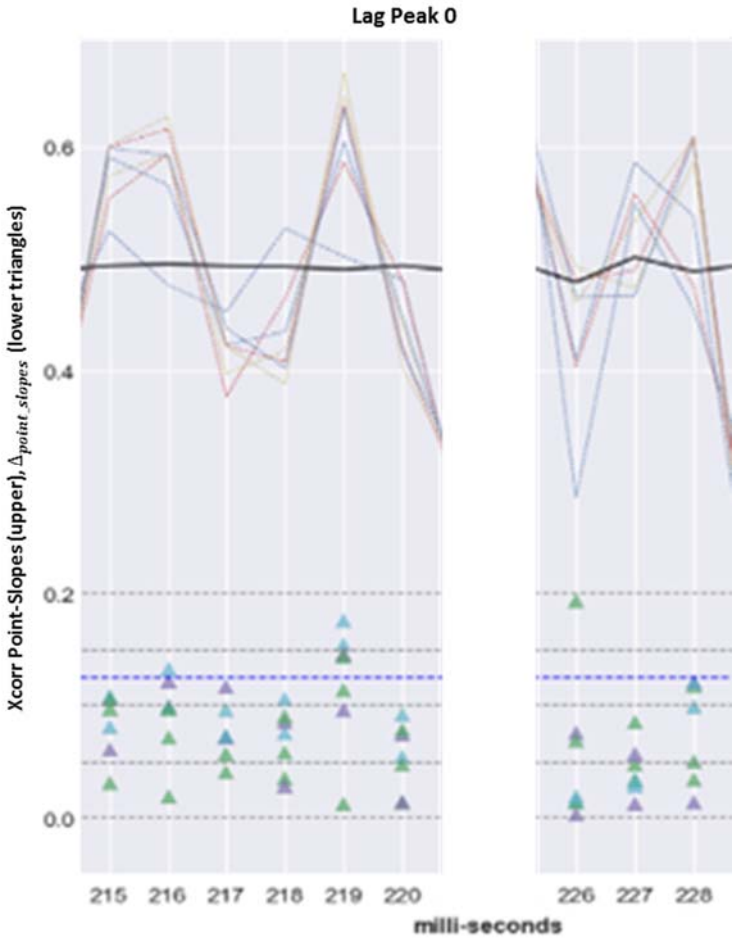

Figure 54 – Time series plot showing 2 windows where the electrode cross-correlation lag peak point-slope differences (triangles) are clustered below a threshold of 0.125 (REF: Figure 47).

CONCLUSION

With an overall project objective to identify pattern-like behaviors upon an Utah Electrode Array dataset, numerous techniques are laid out from the start of this paper that build upon each other to provide an increasingly refined assessment of the 81-electrode time varying voltages. Such techniques span the data analytics spectrum between traditional statistics to other non-traditional approaches that employ a wide variety of plot results as well as improved analytics outcomes concerning the sought for pattern nature of the dataset. The more important of these outputs include: time series, heatmap, and statistical plots, electrode cross-correlation evaluations, and animation results. Using time series voltage plots (Figure 4, Figure 5) and

histogram plots (Figure 16, Figure 19), a general approach and introduction to the basic Utah Electrode Array data was shown and demonstrated as a normal Gaussian distribution for each electrode, permitting a standard statistical approach upon the dataset for further deriving the necessary perspectives. From this, scatter plots were developed to facilitate a succinct method of independent pair-wise electrode voltage visualization. Then, the covariance matrix was established so as to concurrently visualize all 81 pair-wise electrode relationships (Figure 22, Figure 23, and Figure 48). This technique, for the first time, gave visibility to existing pattern-like relationships between all of the dataset electrodes, showing multiple rectangular areas of de-emphasized shading (Figure 48), where these tabulated results are given in Table 1. Additionally, similar dependencies between electrodes using cross-correlation techniques were tested for via calculating 2 electrodes lag peak differences and then measuring against a threshold for the identification of predominate electrode groupings (Figure 51, Figure 52).

A third attempt to investigate electrode patterns was accomplished, where a threshold was again employed and used for the identification of the smallest of lag peak differences upon a time series plot (Figure 43, Figure 44). The visualized calculations show pseudo-synchronous repetition of certain electrodes while others are ruled out according to the specified threshold value. Similarly, the cross-correlation lag peak point-slope differences are plotted (Figure 45, Figure 46, and Figure 47) so as to more deeply understand the specific electrodes effected by the underlying neural network activities. In these plots, particularly the latter and in Figure 54, it is easy to see when a collection of electrode point-slope differences are both approximately the same and also near zero. In this way, certain patterns can potentially be deciphered across a full time series of data, simply by pre-selecting a collection of related electrodes to observe if their results are nearly the same. As Figure 54 offers the best perspective of an observed pattern in this way, notice the 3 right-most time samples clearly showing a single electrode (shown as purple triangles) as it maintains the least point-slope difference consecutively within that window (similarly, observe the left-most window of time samples, shown as green triangles).

Other results are achieved using animated time series plots that are established to show heatmap visualizations of electrode voltages (Figure 8, Figure 9, Figure 32, Figure 33), correlation coefficients (Figure 36, Figure 37), and lag peak similarities (Figure 38, Figure 39). Preceding this,

multiple Look Up Table transformation functions were investigated to achieve the most optimal image contrast and to prevent oversaturation of plot values (Figure 11 thru Figure 13), where a log function was ultimately chosen for both false color and grey color results. Although evidence for the calculated electrode classifications (Figure 51, Figure 52) was visible within the voltage animation results (Figure 53), the derived voltage animations do not sufficiently convey sequentially isolated electrode-to-electrode activities. Thus, an attempt is made to simultaneously observe the 3 cross-correlation lag peaks in a single animation, and although the lag times chosen were pre-selected and therefore not precise for each and every electrode, the outcome did not provide further insight to the desired higher resolution of electrode activity.

FUTURE WORK

The intent of this research study was to visually identify existing neural network patterns from the provided Utah Electrode Array voltage dataset. As discussed throughout this paper, several analytic methods were applied with various degrees of success, however, the moment-to-moment electrode-to-electrode neural patterns across the Utah Array remain elusive. Thus, further endeavor might find progress with extending the previous investigation of lag peak differences plots (Figure 43 and Figure 44) by examining the electrode of least difference at any given time step to determine the closest matching electrode with e01 (the cross-correlation basis). However, this methodology would require that all 81 electrodes are computed via cross-correlation functions such that each lag peak difference result, at each time sample, will provide a single electrode that is nearest in value to the basis electrode, where the final outcome would examine which of the 81 results is nearest to the same result of the prior time sample.

Another future development for this project concerns further analysis of the heatmap animation files created. The capability of computing algorithms today lends itself well to the identification of image patterns using machine learning practices and deep neural network analysis. As such, the already established grey color scale animations are intended for such an input, whereas the false color images are intended for the human reader. Development of machine learning algorithms however, requires the establishing of various data analytics perspectives. As just one example, Figure 55 demonstrates such an outcome where for the first 1,000 time samples of the Utah Array dataset, a Random Forest classifier is evaluated to measure

feature importance of the dataset electrodes. This figure shows that for those electrodes used throughout this paper, electrodes e01, e33, and e44, that only 5 of the 81 electrodes are 'most significant', and furthermore, that the 3 aforementioned electrodes do not appear within this subset of 5 importances. Rather, they are positioned as the 32nd, 71st, and 1st positions respectively. With such a contradictory outcome, it may behoove future researchers to further develop data science and machine learning perspectives into this type of analysis.

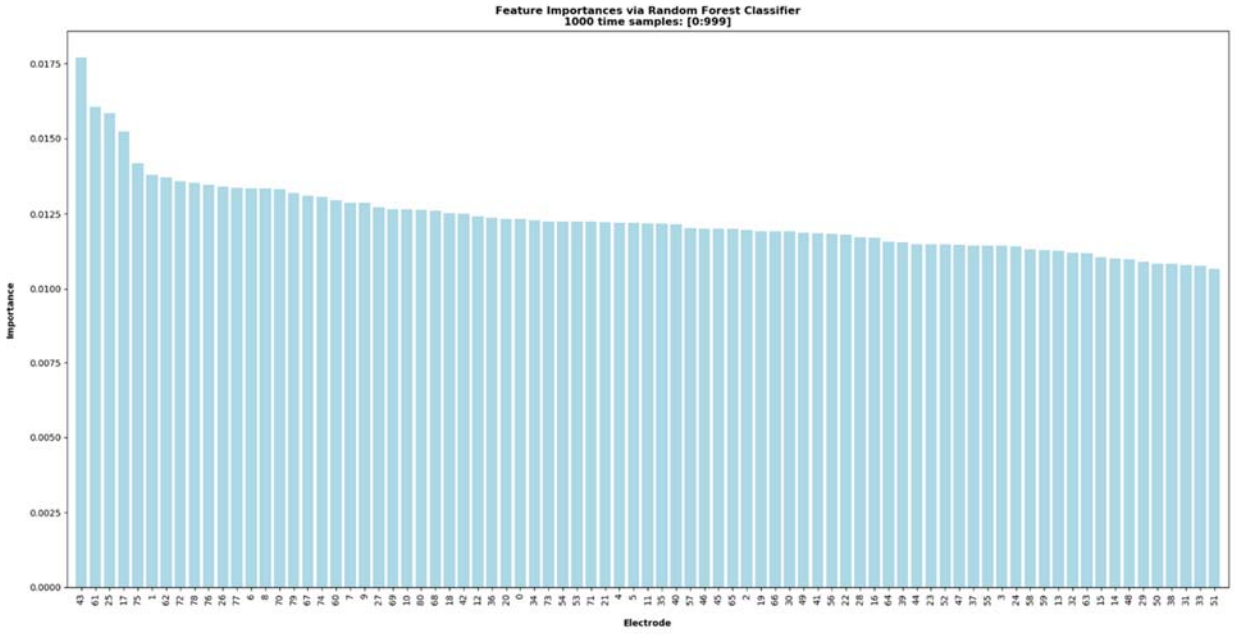

Figure 55 – A plot of electrode feature importances as derived from a Random Forest classifier algorithm.

REFERENCES

[1] D. Edelstein, "Local Field Potentials: Separating Volume Currents from the Local Neural Signals," University of Washington - Dept. of Physics, Seattle, 2012.

[2] "History of Neuroscience," Wikipedia, 27 February 2019. [Online]. Available: https://en.wikipedia.org/wiki/History_of_neuroscience. [Accessed 02 March 2019].

[3] A. Good, "Time Series Analysis of Electrocorticographic Data," University of Washington - Dept. of Physics, Seattle.

[4] D. Winslow, "Time Series Analysis of Brain Data from Patient at Rest," University of Washington - Dept. of Physics, Seattle, 2014.

[5] "Axon," Wikipedia, 21 February 2019. [Online]. Available: https://en.wikipedia.org/wiki/Axon. [Accessed 02 March 2019].

[6] C. T. Nordhausen, P. J. Rousche and R. A. Normann, "Optimizing recording capabilities of the Utah Intracortical Electrode Array," *Brain Research,* no. 637, pp. 27-36, 1994.

[7] E. M. Maynard, C. T. Nordhausen and R. A. Normann, "The Utah Intracortical Electrode Array: a recording structure for potential brain-computer interfaces," *Electroencephalography and clinical Neurophysiology,* vol. 102, pp. 228-239, 1992.

[8] K. A. Ludwig, R. M. Miriani, N. B. Langhals, M. D. Joseph, D. J. Anderson and D. R. Kipke, "Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays," *Journal of Neorophysiology,* vol. 101, no. 3, p. 1679–1689, 2009.

[9] S. H.-F. Syahrull, H. Lakany, R. Ahmad and B. A. Conway, "Comparing Common Average Referencing to Laplacian Referencing in Detecting Imagination and Intention of Movement for Brain Compputer Interface," *MATEC Web of Conferences - ICEESI,* vol. 140, no. 01028, 2017.

[10] "Image Processing II," UC Berkeley, [Online]. Available: http://microscopy.berkeley.edu/courses/dib/sections/03IPII/index.html. [Accessed 17 September 2018].

[11] "Digital Image Processing," Sprawls Educational Foundation, [Online]. Available: http://www.sprawls.org/resources/DIGPROCESS/module.htm. [Accessed 17 September 2018].

[12] "Moving average," wikipedia, 24 01 2019. [Online]. Available: https://en.wikipedia.org/wiki/Moving_average. [Accessed 28 03 2019].

[13] "Standard deviation," wikipedia, 26 03 2019. [Online]. Available: https://en.wikipedia.org/wiki/Standard_deviation. [Accessed 28 03 2019].

[14] "Laerd Statistics - Histograms," Lunde Research Ltd., 2018. [Online]. Available: https://statistics.laerd.com/statistical-guides/understanding-histograms.php. [Accessed 31 03 2019].

[15] D. Q. Nykamp, "The idea of a probability density function," Math Insight, [Online]. Available: https://mathinsight.org/probability_density_function_idea. [Accessed 31 03 2019].

[16] J. L. Rodgers and W. A. Nicewander, "Thirteen Ways to Look at the Correlation Coefficient," *The American Statistician,* vol. 42, no. 1, pp. 59-66, 1988.

[17] A. G. Asuero, A. Sayago and A. G. Gonz´alez, "The Correlation Coefficient: An Overview," *Critical Reviews in Analytical Chemistry,* no. 36, p. 41–59, 2006.

[18] "Linear Regression," Yale University, Dept. of Statistics and Data Science, 1997. [Online]. Available: http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm. [Accessed 06 04 2019].

[19] R. Dalinina, "Introduction to Correlation," ORACLE+DATASCIENCE.COM, 31 01 2017. [Online]. Available: https://www.datascience.com/blog/introduction-to-correlation-learn-data-science-tutorials. [Accessed 05 10 2018].

[20] "Cross-Correlation," wikipedia.com, 18 03 2019. [Online]. Available: https://en.wikipedia.org/wiki/Cross-correlation. [Accessed 08 04 2019].

[21] N. A. Anstey, "Correlation Techniques -- A Review," in *25th European Association of Exploration Geophysicists al Liege*, London, 1964.

[22] "Correlation coefficient review," Khan Academy, [Online]. Available: https://www.khanacademy.org/math/statistics-probability/describing-relationships-quantitative-data/scatterplots-and-correlation/a/correlation-coefficient-review. [Accessed 09 04 2019].

[23] S. Terjung, "Digital Imaging - Data Handling Basic Image Processing," 2018. [Online]. Available: https://www.med.upenn.edu/granato/user_docs/terjung_digital_image_processing.pdf. [Accessed 17 September 2018].