Statistically rigorous calculations do not support Common Input and Long-Term synchronization of motor unit firings

by

Carlo J. De Luca\textsuperscript{1,2,3,4,5,6} and Joshua C. Kline\textsuperscript{1,2}

1) NeuroMuscular Research Center
2) Department of Biomedical Engineering
3) Department of Electrical and Computer Engineering
4) Department of Neurology
5) Department of Physical Therapy

Boston University, Boston MA 02215

and

6) Delsys Inc., Natick, MA 01760

Running title: Novel method measuring motor unit synchronization

Number of text pages: 38, tables: 1, figures: 12

Corresponding Author: Dr. Carlo J. De Luca

NeuroMuscular Research Center
19 Deerfield Street, Boston, MA 02215
Phone (617) 353-9756
Fax (617) 353-5737
cjd@bu.edu
ABSTRACT

Over the past four decades, various methods have been implemented to measure synchronization of motor unit firings. In this work we provide evidence that prior reports of the existence of universal common inputs to all motoneurons and the presence of long-term synchronization are misleading because they did not use sufficiently rigorous statistical tests to detect synchronization. We developed a statistically-based method (SigMax) for computing synchronization and tested it with data from 17,736 motor unit pairs containing 1,035,225 firing instances from the First Dorsal Interosseous and Vastus Lateralis muscles – a data set an order of magnitude greater than that reported in previous studies. Only firing data obtained from surface EMG signal decomposition with greater than 95% accuracy were used in the study. The data were not subjectively selected in any manner. Because of the size of our data set and the statistical rigor inherent to SigMax we have confidence that the synchronization values we calculated provide an improved estimate of physiologically-driven synchronization. When compared to three other commonly used techniques, ours revealed three types of discrepancies that result from failing to use sufficient statistical tests necessary to detect synchronization. 1) On average the z-score method falsely detected synchronization at 16 separate latencies in each motor unit pair. 2) The cumulative sum (cusum) method missed 1 out of every 4 synchronization identifications found by SigMax. 3) The common input assumption method identified synchronization from 100% of motor unit pairs studied. SigMax revealed only 50% of motor unit pairs actually manifested synchronization.

Key words: Motor Unit Firings, Synchronization, Common Input, Synchronization Methods
INTRODUCTION

Over the past four decades, measurements of synchronization of motor unit firing instances have been used to infer the existence of common presynaptic inputs to motoneurons. Typically these measurements have been made from cross-correlation histograms calculated from the firing instances of pairs of motor unit action potential trains (MUAPTs) as described by Perkel et al. (1967b); a work has provided the foundation of all subsequent synchronization measurement techniques. By detecting peaks in the cross-correlation histogram, Sears and Stagg (1976), Nordstrom et al. (1992), McIsaac and Fuglevand (2007), among others, identified firing instances separated by a fixed latency that occurred more often than would be expected by chance. They reported that these peaks indicated the presence of synchronization that resulted from common presynaptic inputs shared by the motoneurons. But, it had been demonstrated previously by Perkel et al. (1967a), Perkel et al. (1967b), and Moore et al. (1970) that cross-correlation peaks could also result from moderately non-stationary firing trains as well as from refractoriness inherent to each neuron. Therefore, the detection of peaks in the cross-correlation histogram alone was known to be insufficient proof that motor unit firing instances are occasionally synchronized as a result of common inputs. For this reason preliminary statistical tests are essential to mitigate the influence of moderate non-stationarities and refractoriness before assessing the degree of synchronization between motor units. However, virtually all synchronization detection methods have not considered these tests and instead have applied assumptions or approximations that remain to be proven by empirical data. Indeed, Perkel et al. (1967b) cautioned that even the most basic statistical assumptions can result in false conclusions from the cross-correlation data, such as false detections of synchronization.
The cumulative sum detection method or cusum technique is one approach commonly used to detect synchronization. This method was originally applied to peristimulus time histograms (PSTHs) by Ellaway (1978) to detect changes in motor unit mean firing rates in response to applied stimuli. Later studies by Connell et al. (1986), Adams et al. (1989), Keen and Fuglevand (2004) among others used the cumulative sum method to identify deviations in the mean value of bin amplitudes from the cross-correlation histogram. These works proposed, but did not prove, that a change in the mean value of the histogram beyond a preset threshold was indicative of synchronization. To our knowledge, no one has documented a correlation between changes in the mean value of the cross-correlation histogram and the relatively high-density regions in the histogram that could result from synchronized motor unit firing instances. Hence, the robustness of the cumulative sum detection method against erroneous detections of synchronization is yet to be established.

After identifying the location of a synchronization peak, many studies rely on the z-score synchronization detection method to compute its statistical significance (Sears and Stagg, 1976; Wiegner and Wierzbicka, 1987; among others). Use of the z-score has been justified by the assumption that bin amplitudes in the cross-correlation histogram of firing instances can be approximated by a normal distribution. Sears and Stagg (1976) and Wiegner and Wierzbicka (1987) advanced this notion by proposing that the Poisson statistics of neuron firing instances reported by Cox and Smith (1954) could be approximated with normal statistics. Although normal statistics may describe the firing instances of some neurons, the actual statistics that describe motoneuron firing instances remain disputed (Lippold et al, 1960; Clamann, 1969; Person and Kudina, 1972; De Luca and Forrest, 1973; among others). Nonetheless, studies by Nordstrom et al. (1990), McIsaac and Fuglevand (2007), Keen et al. (2012) among others, have
not considered empirical reports of non-normal motoneuron firing statistics and instead applied
the z-score to detect synchronization.

Perkel et al. (1967b) illustrated the inadequacy of the z-score method when they noted
that the variance of bin amplitudes in the cross-correlation histogram is greater than that
predicted by normal statistics. Specifically, refractoriness inherent to the motoneuron firing
process induces dependence between adjacent bins in the histogram. As a result, bins with
relatively high numbers of occurrences are more likely to be followed by bins with relatively low
numbers of occurrences. They cautioned that failure to account for the subsequently large
variance of bin amplitudes “can lead to false attributions of dependence to cells that are, in fact,
firing independently.” In spite of this warning, normal approximations of the cross-correlation
histogram bin amplitudes have been used frequently to implement the z-score synchronization
detection method.

Even when a statistically significant synchronization peak is not detected in the cross-
correlation histogram, some studies nevertheless report the degree of synchronization within a
fixed 11 ms time duration centered at 0 ms latency (Semmler and Nordstrom, 1995; Keen and
Fuglevand, 2004; Dartnall et al, 2008; among others). This method is based on the notion that
common inputs produce synchronized firing instances between all pairs of motoneurons. We
refer to this practice as the common input assumption synchronization detection method. It raises
three concerns:

1) It is a subjective approach. The method assumes that synchronization occurs within +/-
5.5 ms centered at a 0 ms latency, even though studies by Kirkwood et al. (1982), Datta
and Stephens (1990), De Luca et al. (1993), Schmied et al. (1993), Semmler and
Nordstrom (1995) among others, have demonstrated that synchronization peaks in the
cross-correlation histogram exist over latencies spanning +/- 20 ms with peak widths ranging from 6 ms to 40 ms. In fact, differences in motoneuron conduction velocities and innervation locations alone could easily produce synchronization latencies as great as 12 ms amongst pairs of motor units from the first dorsal interosseous (FDI) muscle as shown by Andreassen and Arendt-Nielsen (1987), Dengler et al. (1988) and Saitou et al. (2000).

2) The common input assumption synchronization detection method does not apply the mathematically rigorous and empirically tested results of Perkel et al. (1967b) warning that false conjectures of neuron connectivity can arise from computations that lack a sufficient statistical test for dependent firing behavior.

3) The approach occasionally produces negative values of synchronization (Nordstrom et al., 1992), the meaning of which is unclear.

To mitigate these shortcomings, we set out to design an improved approach that is not subject to the above indicated drawbacks. Our statistically based synchronization detection method, which we will refer to as the SigMax detection method, is based on a previous synchronization method developed by De Luca et al. (1993).

Importantly, SigMax does not rely on the assumption that synchronization exists amongst all pairs of motoneurons, and it does not depend on the underlying distribution of the firing instances from each motor unit. The SigMax method is comprised of three tests:

1) **Test for statistically significant non-stationarities** -- Limiting the analysis of synchronization to pairs of stationary MUAPTs is necessary to ensure that synchronization measurements are not biased by moderately non-stationary firing behavior.
2) **Test for statistically significant dependent firing instances** -- The dependence test is applied to pairs of stationary MUAPTs. The test is robust to false detections of synchronization that could result from motoneuron refactoriness.

3) **Test for the most statistically significant peak in the cross-correlation data** -- Only pairs of stationary MUAPTs with dependent firing instances are tested. The detected peak provides the latency, peak width, and magnitude of synchronization between the MUAPTs.

Synchronization was measured using the SigMax detection method and was compared with synchronization calculated by the z-score detection method, the cumulative sum detection method, and the common input assumption detection method. Our analysis demonstrated that these previously used synchronization methods are subject to additional detections, missed observations and disparate estimation of synchronization between MUAPTs.

**METHODS**

*Experimental Design and Protocol*

The experimental design and protocol implemented in this study is described in Kline and De Luca (in review) and will be summarized here. Six healthy subjects, four males and two females, ages ranging from 21 – 23 years, all with no known history of neuromuscular disorders volunteered for the study. Before participating, all subjects read, indicated they understood and signed a consent form approved by the Institutional Review Board at Boston University. All experiments were performed on the first dorsal interosseous (FDI) muscle of the hand and the vastus lateralis (VL) muscle of the lower limb. Isometric force was measured during index finger
abduction and leg extension via load cells. Target trajectories and visual feedback of the isometric contraction force were displayed for the subject on a computer monitor.

The surface EMG (sEMG) signals were recorded with a five-pin sensor previously described in De Luca et al. (2006). The surface sensor was placed on the skin over the center of the muscle belly. Signals from the four pairs of electrodes in the sensor were differentially amplified and filtered with a bandwidth of 20 Hz to 450 Hz. The signals were sampled at 20 kHz and stored in computer memory for off-line data analysis. Before recording data, we measured the maximal voluntary contraction (MVC) force by three brief maximal contractions, each with duration of three seconds separated by a rest period of 3 min. The MVC of greatest value was chosen to normalize the force level of all following contractions for later comparison across subjects. Subjects proceeded to track a series of target trapezoidal trajectories displayed on the computer screen with the output of the force sensor. For the FDI muscle, trajectories increased at a rate of 10% MVC/s, were sustained at 5%, 10%, 15%, 20%, 25% or 30% MVC for 35 s, and were then decreased back to zero at 10% MVC/s. For the VL muscle, trajectories again increased at a rate of 10% MVC/s, were sustained at 20%, 25%, 30%, 35%, 40% or 50% MVC for 35 seconds, and were then decreased back to zero at 10% MVC/s. The recorded force output was band-pass filtered from DC to 450 Hz, digitized at 20 kHz and stored in computer memory for off-line data analysis. At least five minutes of rest were allotted between contractions.

**EMG Signal Decomposition and Error Reduction**

The sEMG signals from four channels of the decomposition EMG (dEMG) sensor were decomposed into their constituent MUAPTs using the dEMG algorithms described by De Luca et al. (2006), substantially improved in Nawab et al. (2010) and independently verified with three
different methods, including direct visual comparison, by Hu et al. (2013a,b,c) and Hu et al.
(2014). The output of the algorithm provided the firing instances of all MUAPTs obtained from
the decomposition. Each firing instance, as measured by the algorithm, was defined by the time
of the greatest absolute value of the action potential.

The occasional errors made by our sEMG signal decomposition algorithm were measured
using the Decompose-Synthesize-Decompose-Compare (DSDC) validation method described by
Nawab et al. (2010). These errors were then mitigated using the error reduction technique
described in the report by Kline and De Luca (in review). In brief, we obtained multiple
independent decomposition estimates, each from the sEMG signal after adding Gaussian-white
noise, equal in RMS to the baseline noise of the sEMG signal. These estimates were then applied
to our error reduction algorithm to produce a new, more probable estimate of the MUAPTs. We
implemented the error reduction procedure using 30 decomposition estimates for each
contraction. Only MUAPTs obtained from decomposition with accuracy greater than 95% were
retained for further analysis.

SigMax Statistical Synchronization Computations:

All computations were performed on a 25 s epoch of sEMG signal recorded during a
constant force isometric contraction. Our SigMax detection method used three statistical tests to
detect and measure synchronization.

Test for statistically significant non-stationarities -- We implemented the now widely
used Kwiatkowski, Philips, Schmidt and Shin (KPSS) test to detect statistically significant non-
stationary MUAPTs (Kwiatkowski et al, 1992). Details of the test are provided in step 1 of the
Appendix. Because moderate non-stationarities could alter the detection and estimation of
dependent firing instances from the cross-correlation data (Moore et al, 1970), it was necessary to perform our synchronization calculations only on stationary MUAPTs.

Test for statistically significant dependent firing instances -- Dependence was assessed from the cumulative distribution of the cross-correlations measured using the recurrence time analysis previously described by Perkel et al. (1967b) and depicted in Figure 1. For each pair of MUAPTs a reference and alternate MAUPT was assigned, with the reference MUAPT having fewer firing instances. Forwards and backwards recurrence times were measured as the latency between each firing instance of the reference MUAPT and forwards and backwards firing instance of the alternate MUAPT, denoted by $t_f$ and $t_b$ respectively. Different orders of recurrence times were computed such that the $i^{th}$ order recurrence times, $t_{f_i}$ and $t_{b_i}$, were measured between the reference firing instance and the $i^{th}$ forwards and $i^{th}$ backwards firing instance of the alternate MUAPT, respectively, as depicted in Figure 1. The value of $i$ ranged from 1 to 5 such that no more than 5 forwards and backwards recurrence times were recorded for each reference MUAPT firing instance.

Figure 1 Near Here

According to basic statistics theory, if two point processes are independent, then they will be uncorrelated. For the specific case of firing instances and recurrence times from pairs of neurons, McFadden (1962) proved that:

if the stationary firing instances from two neurons occur independently

then the recurrence times will be uniformly distributed (uncorrelated).

This is a crucial property of recurrence times fundamental to the study of synchronization. It enables the detection of synchronization that results from dependence between firing instances from pairs of motor units. Using the contrapositive of the proof by McFadden (1962):
if the recurrence times are not uniformly distributed (correlated)

then the stationary firing instances from two neurons do not occur independently;

we tested pairs of stationary MUAPTs for significantly correlated firing instances indicative of dependence. Correlation was evaluated by computing the goodness-of-fit between the empirical cumulative distribution of recurrence times and the predicted uniform cumulative distribution. The specific details of the procedure used for the goodness-of-fit test are provided in step 2 of the Appendix and are summarized here. Recurrence times were divided into different orders at intervals of the mean inter-pulse interval (IPI) of the alternate MUAPT as described by the equation:

\[
Recurrence\ Time\ Interval = i \times IPI_{alt} \pm \frac{IPI_{alt}}{2}; \eq(1)
\]

where \(i\) ranged from -4 to 4. Recurrence times of MUAPTs with independent firing instances were uniformly distributed within each interval of the mean IPI of the alternate MUAPT. Figure 2A shows an example of the empirical cumulative distribution of recurrence times superimposed over the predicted uniform cumulative distribution function. We tested the null hypothesis that the empirical recurrence time data were uniformly distributed using the goodness-of-fit method with the Cramer VonMises test statistic (Cramer, 1928; Von Mises, 1931). Because of its sensitivity to deviations in the mean rather than the variance of the data from uniformity (Stephens, 1974), the Cramer VonMises test was robust to false detections of synchronization that could result from motoneuron refractoriness. The null hypothesis was rejected at the 0.05 significance level corresponding to a Cramer VonMises statistic greater than 0.461 as indicated by Stephens (1970). Recurrence time distributions that significantly deviated from uniformity indicated a statistically significant correlation between the MUAPTs. Significantly correlated
firing instances detected from pairs of stationary MUAPTs indicated the motor units manifested dependent firing instances.

Test for the most statistically significant peak in the cross-correlation histogram -- This test was performed to identify and quantify the amount of synchronization between pairs of stationary MUAPTs with dependent firing instances. Specific details of this test are provided in step 3 of the Appendix and are summarized here. Each of the nine recurrence time intervals from each pair of MUAPTs were separately tested for synchronization by detecting clusters of recurrence times with a density that exceeded what would be expected due to chance. These clusters, or peaks, could occur at different latencies and last for different durations, or peak widths. Our approach iteratively tested all possible latencies and peak widths of the recurrence time data to identify the most statistically significant occurrence of synchronization. Specifically, the peak widths ranged from 1 ms to half of the mean IPI of the alternate MUAPT. For each peak width we detected the latency that produced the greatest number of recurrence time occurrences \((k)\) and computed the statistical significance of the detection. The peak width that produced the number of occurrences with the greatest statistical significance, beyond the 0.05 significance level, indicated a detection of synchronization. An example synchronization detection from the empirical cumulative distribution of recurrence times is shown in Figure 2B. For each detection, the synchronization peak width \((W)\) and latency \((L)\) were recorded for further analysis. The magnitude of synchronization was measured using the synchronization index:

\[
SI = \frac{k_{max} - \bar{k}}{n} \times 100; \; eq(2)
\]
where $k_{max}$ was the actual number of recurrence times detected within the peak of width $W$ and $\bar{k}$ was the average number of recurrence times expected by chance within the peak of width $W$ computed as:

$$\bar{k} = \frac{n}{m}; \quad eq(3)$$

where $n$ is the number of firing instances of the reference MUAPT and $m$ is equal to the mean IPI of the alternate MUAPT divided by the peak width $W$. See Equation 7 in the Appendix. The synchronization index provided the percentage firing instances between two MUAPTs that occurred in excess of chance (De Luca et al, 1993).

Synchronization Measured By Other Methods:

Our SigMax detection method addresses the required statistical considerations for identifying the latency, width and magnitude of synchronization with the greatest statistical significance from each pair of stationary MUAPTs with dependent firing instances. We compared synchronization results obtained from SigMax with those obtained by three other previously published methods that do not incorporate statistical tests necessary to adequately measure synchronization. This comparison was performed on the same set of stationary MUAPTs evaluated using the SigMax method. Synchronization was quantified using the synchronization index in Equation 2 for all detection methods tested.

Z-score synchronization detection method -- We followed the practice of assuming a normal distribution of bin amplitudes in the cross-correlation histogram to compute the z-score significance threshold. The baseline mean amplitude of the histogram was measured over a +/- 200 ms, region of the histogram, excluding +/- 20 ms region centered around 0 ms latency, as prescribed by previous methodologies used by Sears and Stagg, 1976; Schmied et al, 1993; Keen
and Fuglevand, 2004; among others. We then identified peaks in the histogram that exceeded the 0.05 normal significance threshold corresponding to a z-score > 1.96. The number of significant peaks detected and the corresponding synchronization index, peak width and latency were compared to the synchronization measured for the same MUAPT pairs using our SigMax detection method.

We also tested the underlying assumption used to justify the appropriateness of the z-score for detecting synchronization. As stated in the introduction, the z-score method is based on the assumption that motor unit firing instances are normally distributed. Previously used distribution tests relied on Chi-squared or Komolgorov-Smirnov (KS) test statistics to identify the deviations of empirical data from the expected distribution (Clamann, 1969; De Luca and Forrest, 1973; among others). While these methods were the most practical for the time of their use, later work by Stephens (1974) showed that the Chi-squared and KS tests have lower statistical power than other goodness-of-fit test statistics. For example in the specific case of normal hypothesis testing, D’Agostino et al. (1990) showed that these tests are prone to missed detections of deviations from normality. In contrast, the D’Agostino-Pearson omnibus test (D’Agostino and Pearson, 1973) overcomes these drawbacks not only because it is more sensitive to non-normal statistics, but because it quantifies the nature of deviation from normality using skewness and kurtosis statistics. Therefore, we implemented the D’Agostino-Pearson omnibus test to determine if the motor unit IPIs significantly deviated from a normal distribution. The null hypothesis of normality was rejected at the 0.05 significance level.

Cumulative sum synchronization detection method -- This method detects a synchronization peak using the cumulative sum of bin amplitudes from the cross-correlation histogram. We applied the cumulative sum detection method by first measuring the baseline
mean of the cross-correlation histogram within +/- 200 ms, excluding the +/- 20 ms centered around 0 ms latency. The cumulative sum data were then computed as the running sum of the difference between the baseline mean and the amplitude of each bin in the cross-correlation histogram. For additional details see Ellaway (1978). The cumulative sum data were normalized by the difference between the maximum and minimum values. Locations where the normalized cumulative sum data crossed the 0.1 and 0.9 thresholds indicated the positive and negative boundaries of a synchronization peak respectively (Schmied et al, 1993; Keen and Fuglevand, 2004; among others). The synchronization peak was calculated between the boundaries of the cumulative sum data and was compared with the synchronization peak obtained from our SigMax detection method.

Common input assumption synchronization detection method -- This method computes the magnitude of synchronization from a fixed region of the cross-correlation histogram regardless of the statistical significance computed for the detected amount of synchronization. According to this approach, synchronization was measured from a fixed 11 ms region of the cross-correlation histogram centered at 0 ms latency (Semmler and Nordstrom, 1995; McIsaac and Fuglevand, 2007; Keen et al, 2012). Results from the common input assumption synchronization detection method were compared to synchronization results obtained using our SigMax detection method.

RESULTS

SigMax Synchronization Results

All MUAPTs were obtained using 30 iterations of the error reduction algorithm performed on 144 recorded sEMG signals; a subset of 36 of these signals were used by Kline and
De Luca (in review) to thoroughly test the error reduction process. In total, 2,287 MUAPT s were obtained with accuracies above 95%; 894 were from 72 FDI contractions ranging from 5 to 30% MVC and 1,393 were from 72 VL contractions ranging from 20 to 50% MVC. In total statistically significant non-stationarities were detected in 100 or 11.2% of MUAPTs from FDI contractions and 187 or 13.4% of MUAPTs from the VL contractions. The remaining stationary data consisted of 794 MUAPTs with 333,633 firing instances from the FDI and 1,206 MUAPTs with 701,592 firing instances from the VL.

Of the stationary data, the IPIs from 98.6% of MUAPTs from the FDI and 99.8% of MUAPTs from the VL significantly deviated from a normal distribution. Examples of non-normal IPI histograms from several MUAPTs are shown in Figure 3. Normally distributed IPIs would manifest skewness equal to 0 and kurtosis equal to 3 (D’Agostino et al, 1990). But according the D’Agostino-Pearson omnibus test, the IPIs of all six MUAPTs in Figure 3 manifested skewness and kurtosis values that significantly deviated from normality (p<0.0001). On average, IPIs from all MUAPTs tested were positively skewed and had relatively greater occurrences in the distribution tails than expected by normal statistics (Table 1).

In total, 6,453 and 11,283 pairs of stationary MUAPTs were tested for synchronization from FDI and VL contractions respectively. Figure 4 summarizes the results. Our SigMax detection method found statistically significant synchronization in 42.0% of FDI MUAPT pairs and 54.8% of VL MUAPT pairs. The average synchronization index was 19.8 and 16.9, the synchronization peak width averaged 25.8 ms and 18.5 ms and the synchronization latency averaged -0.1 ms and -0.3 ms in FDI and VL data, respectively.

Table 1, Figure 3 and Figure 4 Near Here
Effects of Decomposition Errors on Calculations of Synchronization -- To illustrate the importance of mitigating decomposition errors before measuring synchronization, we compared results from our SigMax method applied to pairs of MUAPTs obtained with and without the error reduction algorithm. Synchronization detections were compared between the same pairs of MUAPTs in both muscles. According to the results shown in Figure 5, unmitigated decomposition errors resulted in 11.2% or 1,268 additional detections of synchronization and 22.1% or 2,498 missed detections of synchronization. For those pairs of MUAPTs in which synchronization was correctly identified, on average decomposition errors produced a synchronization index that differed by of 21.5%, a peak width that differed by of 25.1% and a peak latency that differed by more than +/-100% (Figure 5).

Figure 5 Near Here

Comparison of the Z-Score Synchronization Detection Method with SigMax -- Figure 6 presents an example of z-score synchronization detections from the cross-correlation histogram of recurrence times for one pair of stationary MUAPTs. The average bin amplitude of the histogram is indicated by the horizontal dashed line and the 1.96 z-score statistical significance threshold is shown by the horizontal solid line. In total, 12 peaks manifested an average value above the normal significance threshold, indicating 12 separate detections of synchronization from the z-score method. For the same data, the SigMax detection method identified only 1 statistically significant synchronization peak. When applied to all the stationary MUAPTs the z-score synchronization detection method identified 104,662 synchronization peaks from 6,453 pairs of FDI MUAPTs and 177,894 synchronization peaks from 11,283 pairs of VL MUAPTs (Figure 7). On average the z-score detection method found 16 synchronization peaks from each
pair of MUAPTs. This compares to a single synchronization detection found by SigMax in only 50% of the paired MUAPTs studied.

We evaluated the capabilities of the z-score detection method to estimate synchronization in the most centrally located peak in each cross-correlation histogram. The average values and 95% confidence intervals of the synchronization index, peak width and latency (bars with diagonal lines in Figure 4) were calculated for the same set of stationary MUAPTs evaluated using the SigMax method. Relative to the results from the SigMax method, on average the z-score method produced a synchronization index that differed by -71.9%, a synchronization peak width that differed by -85.8% and a synchronization latency that differed beyond +/-100% (bars with diagonal lines in Figure 7).

Figures 5 and 6 Near Here

Comparison of the Cumulative Sum Synchronization Detection Method with SigMax --

Two examples of the cumulative sum synchronization detection method are shown in Figure 8. The cross-correlation histogram in Figure 8A provides an example of recurrence time data from the first recurrence interval of MUAPT pair #1. The corresponding normalized cumulative sum for the histogram is plotted in Figure 8B. The 0.1 and 0.9 confidence intervals, typically used in the cumulative sum detection method (Schmied et al, 1993; Keen and Fuglevand, 2004; among others), are illustrated by the horizontal dashed lines. The cumulative sum data crossed the 0.9 confidence interval near 19 ms. However, the 0.1 confidence interval was not crossed within the first recurrence interval shown. The same data analyzed using our SigMax detection method revealed a statistically significant peak centered near 10 ms that spanned 14 ms in width (Figure 8C). Because the cumulative sum method was unable to identify a synchronization peak within the first recurrence interval, MUAPT pair #1 was marked as a missed detection.
The cross-correlation histogram measured from MUAPT pair #2 is shown in Figure 8D, with the corresponding normalized cumulative sum displayed in Figure 8E. A synchronization peak was detected by the cumulative sum data between -9 ms and 4 ms. Relative to the 22.7 ms wide synchronization peak found by our SigMax method (Figure 8F) the cumulative sum detection method produced a synchronization peak width that differed by 9.7 ms or 42.7%.

When we applied the cumulative sum method to the same set of stationary MUAPTs evaluated using the SigMax method, synchronization peaks were identified from 99.8% of the paired MUAPTs tested. The peak widths of the cumulative sum detections spanned +/-200 ms of the cross-correlation histogram tested. However, to better evaluate differences between the cumulative sum detection method and our SigMax method, we limited the analysis of cumulative sum peaks to only those detected within the first interval of recurrence times. For these data, the cumulative sum method detected synchronization in 43.3% of paired MUAPTs from the FDI and 36.9% of paired MUAPTs from the VL (bars with dots in Figure 4). Relative to synchronization detected by the SigMax method, the cumulative sum method produced 6.2% or 549 additional detections of synchronization and 27.5% or 2,447 missed detections of synchronization. When comparing the detections made by both methods, the greatest discrepancies in the synchronization statistics were observed for the peak widths, 95% of which differed within the range of -67.6% to 101% (bars with dots in Figure 7).

Comparison of the Common Input Assumption Synchronization Detection Method with SigMax -- Two sample synchronization peaks are shown in Figure 9. For one pair of MUAPTs, the common input assumption method assumed that the synchronization peak was located at 0 ms latency with a peak width of 11 ms, eliciting a synchronization index of -8.14 (Figure 9A).
For the same pair of MUAPTs our SigMax detection method indicated that no statistically significant synchronization peak was present.

The cross-correlation histogram measured from the recurrence times of a second pair of MUAPTs is shown in Figure 9B. The common input assumption detection method assumed the synchronization peak was located at 0 ms latency with a peak width of 11 ms. This assumed detection yielded a synchronization index of 0.536. However, our SigMax detection method found the most statistically significant peak in the data centered near 10.4 ms latency with a peak width of 6.78 ms and synchronization index of 5.88.

The common input assumption method detected synchronization peaks in 100% of the pairs of stationary MUAPTs; whereas, the SigMax method detected synchronization in only 50% of stationary MUAPT pairs. When compared to the SigMax method, on average the synchronization index differed by -35.0%, the synchronization peak width differed by -38.8% and the synchronization latency differed by -100% (bars with diamonds in Figure 7). But, most peculiarly, the common input assumption method produced negative values of synchronization in 2,339 or 13.2% of paired MUAPTs (bars with diamonds in Figure 4).

Long-term Synchronization

The proportion paired MUAPTs detected with synchronization by our SigMax method at each recurrence interval are shown in Figure 10A and 10C for FDI and VL data respectively. The first recurrence interval manifested the greatest percentage of pairs of MUAPTs with synchronization. Relatively higher order recurrence intervals had progressively fewer synchronized MUAPT pairs.
Perkel et al. (1967b) observed that peaks in relatively lower-orders of recurrence times often produced harmonic peaks in relatively higher-orders of recurrence times. Their analysis demonstrated these harmonic peaks are merely statistical artifacts of recurrence time analysis. Such occurrences do not necessarily indicate unique occurrences of synchronization. The gray shaded bars in Figure 10A and 10C illustrate the percentage of pairs of MUAPTs with synchronization that did not result from harmonics. All recurrence time intervals beyond the first interval manifested less than 4% of paired MUAPTs with synchronization. Because our SigMax detection method employed a minimum significance threshold of 0.05, as many as 5% of MUAPT pairs could manifest significant peaks as a consequence of random variance in the recurrence time data that were not indicative of synchronization. Therefore, all recurrence intervals with fewer than 5% of synchronized MUAPT pairs were considered statistically insignificant detections. The remaining statistically significant detections of synchronization occurred within the first recurrence time-interval. The final distributions of the latencies of these synchronization detections are shown in Figure 10B and 10D. Overall, 95% of the synchronization latency data ranged from -7.4 to 7.2 ms in the FDI and from -6.0 to 5.5 ms in the VL.

DISCUSSION

This study revealed two physiological findings:

1) Our analysis did not support the previously reported assumption that common inputs cause synchronization amongst all motoneurons in a pool.

2) Long-term synchronization should not be interpreted as representing a physiological event.
The Common Input Assumption

We found synchronous and dependent firing behavior in only 50% of the motor unit pairs studied. Our result differs from that predicted by the notion that common inputs cause synchronization amongst all pairs of motoneurons in the pool of a given muscle. Yet, our finding is consistent with the original work of Sears and Stagg (1976) who concluded that the common input notion “does not mean that each contributing motoneuron necessarily has common presynaptic inputs with every other motoneuron of the pool.” We first reported our doubts concerning the common input assumption in De Luca et al (1993). More recent studies by Keen and Fuglevand (2004), Hockensmith et al. (2005), Dartnall et al (2008), have also reported that only a fraction of motor unit pairs manifest statistically significant synchronization. However, these later studies did not consider the statistical significance of their results and continued to support the notion of common inputs between all pairs of motoneurons.

Consider the underlying requirements to prove the existence of the common input assumption. Dependence can only be proven between pairs of stationary MUAPTs using methods capable of detecting significantly correlated firing instances that are robust to false detections from motoneuron refractoriness. According to basic statistical theory:

\[
\text{if the firing instances of two motor units are independent}
\]

\[
\text{then the stationary firing instances are uncorrelated.}
\]

Then according to the contrapositive:

\[
\text{if the stationary firing instances are correlated}
\]

\[
\text{then the firing instances of two motor units are not independent.}
\]

Dependence between firing instances of paired motor units can only be proven by first establishing that the firing instances manifest a statistically significant correlation. Although this
principle is fundamental to measuring synchronization, it has not been taken into account in the majority of previous studies. Semmler and Nordstrom et al (1995), McIsaac and Fuglevand (2007), Keen et al. (2012), among others proposed that the magnitude of synchronization quantified the strength of common inputs received by all pairs of motoneurons. Their approach assumed dependence of all motoneuron pairs on common inputs without proving that the motor unit firing instances were correlated. Yet the inverse of the conditional relationship above indicates that:

\[ \text{if the stationary firing instances are uncorrelated,} \]
\[ \text{then the firing instances of two motor units may or may not be dependent.} \]

Therefore, measurements of synchronization from paired motor units with uncorrelated firing instances do not prove that pairs of motoneurons receive common inputs.

Even if Sears and Stagg (1976), Kirkwood & Sears (1978) and Nordstrom et al (1992) observed significantly correlated firing instances between some pairs of motor units, their observations still would not prove that common inputs cause synchronization. Consider the following simple example. An observer records a person walking on the beach approximately every twelve hours. The observer also records that the tide of the ocean is highest every time the person walks on the beach. After repeatedly making similar observations over several days, the observer notices a correlation between the person walking on the beach and the high tide. Although the correlation in the data indicates some degree of dependence between the person walking and the high tide, it does not prove that the person causes the high tide. Similarly, correlated firing instances indicative of synchronization between pairs of stationary MUAPTs do not prove that synchronized firing instances are caused by common inputs to the motoneurons.
Long-term Synchronization

Long-term synchronization with latencies beyond +/-20 ms has been reported by Kirkwood et al. (1982), Datta and Stephens (1990), De Luca et al. (1993), Schmied et al. (1993), Semmler and Nordstrom (1995), among others. Using the SigMax detection method, we found that these occurrences are not likely to be a physiological event, but are an artifact of false detections produced by two factors inherent to previous synchronization detection methods.

The first factor is the use of an insufficient and relatively low significance threshold to detect synchronization peaks. This effect may be seen in Figure 6 for the z-score method where an overwhelming number of synchronous peaks were detected at long-latencies. Similar detections of long-term synchronization were obtained in 18% of the motor unit pairs using a binomial significance threshold as done by De Luca et al. (1993). Comparison of their results to those presented in this study suggests that their binomial significance threshold underestimated the actual statistical significance of synchronization detections.

The second factor is the presence of harmonics in the cross-correlation histogram (Figure 10). Specifically, when synchronization is observed amongst first-order recurrence times, there is a greater likelihood that synchronization will also be observed at higher-orders of recurrence times. The presence of these harmonics has been well documented in the work of Perkel et al. (1967b), but has not been considered by most studies of synchronization. (For examples see Kirkwood et al, 1982; Bremner et al, 1991; Semmler et al, 2002; among others). The error can be avoided by calculating synchronization exclusively from first-order recurrence times using an objectively derived and statistically reasoned method to detect statistically significant occurrences of synchronization from stationary MUAPTs with dependent firing instances.
Quality of data

In addition to using a more robust and statistically more rigorous approach to measuring synchronization, we had a vetted data set representing physiologically meaningful information relevant to manifestations of synchronization during isometric contractions.

1) The data had a known high accuracy level. The decomposition algorithm used in this study has been extensively validated by De Luca et al. (2006), Nawab et al. (2010), De Luca and Contessa (2012). Independent verification using three different methods by Hu et al (2013a,b,c) and Hu et al (2014) has confirmed that our dEMG algorithms can yield firing instances of MUAPTs having an average accuracy of 95%. Furthermore, we applied the error reduction algorithm described in Kline and De Luca (in review) to reduce decomposition errors and improve the estimate of the identification and the location of the firing instances. As may be seen in Figure 5, our analysis showed that unmitigated decomposition errors result in false identifications, missed detections and incorrect estimation of synchronization. Other studies of synchronization have reported no measure of the decomposition accuracy. Instead, to address potential decomposition errors, firing instances located amongst superpositions of other action potentials were discarded (Nordstrom et al, 1990; Nordstrom et al, 1992; Semmler and Nordstrom, 1995). It is in these locations that different MUAPs occur at or near the same time indicating the firing instances of different motor units are most likely to be synchronized. This tailoring practice would likely contribute to the lesser synchronization levels reported previously.

2) The data set was assembled objectively. MUAPTs were not subjectively culled or altered in any manner. Instead, we performed our synchronization analysis on the entire set of MUAPTs obtained from sEMG decomposition within the accuracy bounds we report.
3) Combined with our technology for recording and decomposing sEMG signals, our experiments were designed to study synchronization in natural voluntary contractions where the subjects were instructed to maintain a constant force. This paradigm is similar to that used by De Luca et al (1993), Contessa et al (2009) and Defreitas et al (2013). Due to technical limitations, other studies were constrained to study MUAPTs during artificial contractions, where the subjects were typically instructed to manipulate their contractions so as to maintain fixed motor unit firing rates around 8-10 pps (Datta and Stephens, 1990; Nordstrom et al, 1992; Semmler et al, 1997; Keen and Fuglevand, 2004; Dartnall et al, 2011; among others). Force levels were rarely reported, but from the description of the tests they seemed to be below 5% MVC. In our experiments data were obtained from more natural contractions whose force levels ranged from 5% to 50% MVC.

Evolution of the SigMax Detection Method

Our approach for measuring synchronization of firing instances between pairs of MUAPTs considered the consequences of incorrect statistical assumptions cautioned by Perkel et al. (1967b). We addressed their warnings by designing a synchronization method that detected the most statistically significant incidence of synchronization from pairs of stationary MUAPTs with dependent firing instances. The tests for stationarity and dependence are fundamental precursory measures for calculating the degree of synchronization between motor units. For the past four decades, all of the methods that have been developed to measure synchronization have not applied these considerations. Some proposed that the motor unit firing instances had a Gaussian distribution, but never tested their assumptions. Others established values of the latency and peak width of synchronization without testing statistical significance. And almost all
methods did not test for statistical dependence between the pairs of MUAPTs but were used nonetheless to substantiate the notion that motoneurons depend on common inputs.

The SigMax detection method overcomes the shortcomings of other approaches. It is an extension of the method we described in De Luca et al. (1993). Like the earlier report it employs binomial statistics to detect synchronization. Our earlier work, although being more statistically rigorous than other methods had a drawback. The method underestimated the probability of detecting regions of statistically significant density in the cross-correlation histogram. To overcome the short-coming of our earlier work, instead of computing the 95% binomial significance threshold that each bin has a given number of occurrences, we computed the 95% binomial significance threshold that any bin has a given number of occurrences. Mathematically, this difference is illustrated by comparing the single binomial probability equation provided in De Luca et al. (1993) to the more detailed series of equations required for our approach as detailed in step 3 of the Appendix. Probabilistically the difference becomes visible when considering a simple experiment of rolling dice. The probability of rolling three number 3’s from a six sided die after \( n \) rolls is lower than the probability of rolling three of any kind from a six sided die after \( n \) rolls. As a result the 95% binomial significance threshold of rolling three number 3’s is relatively lower than the 95% binomial significance threshold of rolling three of any kind. This accounts for the lower significance threshold used by De Luca et al. (1993). In contrast, our significance threshold is relatively higher and therefore less susceptible to the false detections. Specifically, the following improvements were made:

1) Our SigMax method does not require any a-priori assumptions of the underlying motor unit firing statistics. This contrasts with the z-score synchronization detection method that depends on a normal distribution of bin amplitudes in the cross-correlation histogram –
Novel method measuring motor unit synchronization

an assumption that results in erroneous detections of synchronization. Instead, SigMax identifies the most statistically significant detection of synchronization, regardless of the statistics that describe motor unit firing instances.

2) We implemented a test to identify MUAPTs that contained statistically significant non-stationarities – a fundamental preliminary measure that must precede any synchronization analysis. Only stationary MUAPTs were used for the analysis of synchronization. Data from Moore et al. (1970) demonstrated that moderate non-stationarities hinder the detection and estimation of dependent firing behavior between pairs of MUAPTs.

3) Pairs of stationary MUAPTs were directly tested for significantly dependent firing instances. Our goodness-of-fit test evaluated by the Cramer-Von Mises test statistic ensured that all measurements of synchronization made between stationary MUAPTs were proven to have dependent firing instances. All other reported approaches used to measure synchronization did not apply tests for dependent motor unit firing behavior.

4) We measured synchronization from the cumulative distribution function of cross-correlation data, rather than from the cross-correlation histogram of the data. In so doing, we avoided the false detections of synchronization that could occur from relatively high fluctuations of bin amplitudes as a result of the arbitrary nature of histogram binning. The use of the cumulative distribution function also allowed us to avoid false detections of synchronization from relatively high fluctuations in bin amplitudes that typically result from motoneuron refractoriness (Perkel et al, 1967b).

Consequence of the Z-Score Synchronization Detection Method
The z-score detection method relies on the assumption that the distribution of the IPIs is Gaussian. Contrary to previous reports by Buchthal et al. (1954), Clamann (1969), and Andreassen and Rosenfalck (1980), the normality of the IPI distribution has been questioned before by Lippold et al. (1960), Person and Kudina (1972), De Luca and Forrest (1973), among others. The overwhelmingly conclusive results from the motor unit IPI normality test performed in this study demonstrate that IPIs of MUAPTs in EMG signals obtained from natural (unconstrained) isometric contractions are not normally distributed. We are confident in this observation because: 1) the size of the data from individual contractions was considerably greater (more than an order of magnitude) than that used in previous studies; 2) we used the D’Agostino-Pearson omnibus test for normality which has been widely accepted as a superior test than those available to the earlier reports; and 3) the number of observations used for each normality test was previously tested by D’Agostino et al. (1990) and found to be sufficient to reveal statistics underlying motor unit firing instances. Some previous reports of the z-score method (Sears and Stagg, 1976; Nordstrom et al, 1992; Keen et al, 2012) may have used data sets with greater number of firing instances, but the D’Agostino-Pearson omnibus test indicates that such large data sets are not required. In fact, the presumed need for large data sets may have cajoled investigators to collect data under the artificial (constrained) contraction paradigm discussed previously.

Consequences of the Cumulative Sum Synchronization Detection Method

Relative to our SigMax method the cumulative sum method produced two types of discrepancies when estimating the synchronization peak: 1) approximately 1 in every 4 synchronization peak detections are missed (Figure 7A, 7B); and 2) correctly identified peaks
manifest peak width discrepancies that range beyond +/-50% (Figure 7C – 7H). Although synchronization detections from the cumulative sum method are subject to relatively large estimation discrepancies, the average synchronization values are relatively similar to those measured using SigMax. This result indicates that comparisons of average values may not necessarily reveal the discrepancies produced by a given method. Overall, the discrepancies reported in Figure 7 demonstrate that changes in the cumulative sum from the mean value of the cross-correlation histogram do not guarantee adequate detections of synchronization.

**Consequence of the Common Input Assumption Synchronization Detection Method**

The common input assumption synchronization detection method established *de facto* that the latency of synchronized firing instances occurs at 0 ms spanning 11 ms in peak width. Our analysis has demonstrated that this assumption is factually incorrect and resulted in the detection of negative values of synchronization from 13.2% of pairs of FDI and VL MUAPTs. Nordstrom et al. (1992) also reported negative synchronization. The physiological interpretation of negative synchronization has never been explained and remains a confounding detail that poses a physiological conundrum.

Overall, our analysis motivates the need for robust statistical methods when measuring synchronization. Failure to apply necessary fundamental tests, such as those for stationarity or dependence, lead to the discrepancies between synchronization we observed and that reported in previous studies. It is only from empirical data obtained by factually substantiated statistical measures that physiological mechanisms of motor unit control can be revealed properly. As such the validity of the notion that common inputs cause synchronization of motor unit firing instances awaits experimental verification.
We are grateful to Dr. RA D’Agostino for guidance in development of the statistical procedures used for this analysis. We thank the subjects who painstakingly participated in the experiments.

This work was supported in part by two grants from NIH [HD05011/HD/NICHD and NS077526-01/NS/NINDS], and funds from Delsys Inc.

Carlo J. De Luca is the President and CEO of Delsys Inc, the company that developed the sEMG decomposition technology.

CJ De Luca: Conception and design of research; interpreted results of experiments; drafted manuscript; edited and revised manuscript; approved final version of manuscript.

JC Kline Conception and design of research; performed experiments; analyzed data; interpreted results of experiments; prepared figures; drafted manuscript; edited and revised manuscript; approved final version of manuscript.
APPENDIX – SigMax Synchronization Detection Method

Step 1: Test for statistically significant non-stationarities -- Methods that were previously used to detect non-stationary MUAPTs have relied on qualitative analyses such as observing scatter plots of successive IPIs (Clamann, 1969; Masland et al, 1969; Person and Kudina, 1972; among others). Although innovative at their time of use, these techniques limit the detection of non-stationarities to specific orders of IPIs. To overcome this shortcoming, we implemented the now widely used Kwiatkowski, Philips, Schmidt and Shin (KPSS) test to detect statistically significant non-stationary MUAPTs (Kwiatkowski et al, 1992). According to DeJong et al. (1992) and Diebold and Rudebusch (1991) the KPSS test has greater statistical power for detecting non-stationary data than the standard unit root tests presented by Dickey and Fuller (1979). For each train of IPIs we tested the null hypothesis that the motor unit firing instances separated by different lag times were stationary. We implemented the test with $T^{1/2}$ different lag times; where $T$ is the sample size of IPIs for each MUAPT, as recommended by Andrews (1991) and Kwiatkowski et al. (1992). The null hypothesis was rejected at the 0.05 significance level, corresponding to a KPSS test statistic of 0.463. All MUAPTs that produced a KPSS test statistic greater than 0.463 contained statistically significant non-stationary firing instances. Only stationary MUAPTs were tested for synchronization.

The IPIs from two example MUAPTs are shown in Figure 11. Specifically, in Figure 11A not only was no visible trend apparent in the data, but the KPSS test statistic was 0.0643, well below the 0.463 detection threshold for non-stationary MUAPTs. In Figure 11B, the IPIs of a moderately non-stationary MUAPT are shown. Although only a slight positive trend was visible in the IPI data as a function of time, the KPSS test statistic was 1.19, indicating the MUAPT contained statistically significant non-stationary firing instances.
Step 2: Test for statistically significant dependent firing instances -- We assessed
dependence between firing instances from pairs of stationary MUAPTs by identifying
statistically significant correlations in the recurrence time data using a goodness-of-fit test
between the empirical cumulative distribution function of recurrence times and the predicted
uniform cumulative distribution function. Further details of this and other goodness-of-fit tests
are provided in D’Agostino and Stephens (1986). Specifically, we tested the null hypothesis that
if the firing instances of two motor units are independent, then their recurrence times will be
uniformly distributed (McFadden, 1962). Mathematically this was represented by the hypothesis:

\[ H_0: t_1, t_2, \ldots, t_n \text{ comes from } U(t), \]

that \( n \) recurrence times, \( t_i \), come from a uniform distribution. By detecting deviations of the
empirical cumulative distribution function of recurrence times, \( F(t) \), from the predicted uniform
cumulative distribution function, \( U(t) \), we could reject the null hypothesis and identify motor
units with dependent firing instances. We evaluated the goodness-of-fit between the empirical
and uniform cumulative distributions using the Cramer-Von Mises test statistic, \( W_n^2 \), first
presented by Cramer (1928) and Von Mises (1931), then later improved upon by Smirnov (1936,
1937) as:

\[
W_n^2 = n \int_{-\infty}^{\infty} [F(t) - U(t)]^2 dU(t); \quad eq(4).
\]

We chose the Cramer-Von Mises test statistic, \( W_n^2 \), as opposed to the Chi-Squared or
Komogorov-Smirnov statistics because the Cramer-Von Mises test statistic tends to have greater
statistical power (Stephens, 1974). To compute the Cramer-Von Mises goodness-of-fit test we
arranged the recurrence times, \( t_i \), in ascending order such that:

\[ t_1 \leq t_2 \leq \cdots \leq t_n. \]
We then performed a probability integral transformation (Fisher, 1930) of the recurrence time data to obtain the sequence of values $u_i$ such that:

$$u_i = U(t_i); \text{ for } i = 1 \ldots n.$$  

These data were used to compute the Cramer-Von Mises test statistic specific for the uniform distribution goodness-of-fit test derived by Pearson and Stephens (1962) and Stephens and Magg (1968) as:

$$W_n^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left[ u_i - \frac{(2i - 1)}{2n} \right]^2; \text{ eq}(5).$$

Following the statistical analysis of Stephens (1970) we adjusted $W_n^2$ to produce a modified Cramer-Von Mises test statistic, $W^*$, as:

$$W^* = \left( W_n^2 - 0.4/n - 0.6/n^2 \right) \left( 1.0 + 1.0/n \right); \text{ eq}(6).$$

Using Table 1 in Stephens (1970), the modified Cramer-Von Mises test statistic, $W^*$, was compared with the derived percentage points of $W^*$ to determine the statistical significance of the goodness-of-fit. We rejected the null hypothesis, $H_0$, at the 0.05 significance level demonstrated by the conditional:

\[ \text{if } W^* > 0.461 \text{ then } \rightarrow \text{Reject } H_0. \]

Rejection of the null hypothesis, $H_0$, established the existence of statistically significant correlations indicating dependence between the firing instances of the stationary MUAPTs.

**Step 3: Test for the most statistically significant peak in the cross-correlation histogram** - We identified the most statistically significant occurrences of synchronization from pairs of stationary MUAPTs with dependent firing instances. Figure 12 illustrates a pictorial demonstration of our algorithmic approach. For each empirical cumulative distribution of recurrence times we first selected a temporal window of width $W_f$ such that: 
\[ W_j = \frac{IPI_{Alt}}{m_j}; \quad \text{eq}(7) \]

where \( m_j \) ranged from 2 to the closest integer greater than or equal to the mean IPI of the alternate MUAPT (\( IPI_{Alt} \)). The window width \( W_j \) was chosen such that the recurrence time data over the given interval could be entirely sectioned into \( m_j \) equal sections of width \( W_j \). For each window width, the number of \( k \) recurrence times within the window was counted at different latencies, \( L \) (Figure 12). The latency, \( L \), was incrementally increased by 0.1 ms. After computing the amplitude \( k \) within the window of width \( W_j \) at all possible latencies, we identified the latency, \( L \), that contained the maximum amplitude of recurrence times, denoted \( k_{max} \).

![Figure 12 Near Here](image)

We computed the probability of finding at least \( k_{max} \) occurrences within the interval of width \( W_j \). The probability that any of \( m_j \) equally spaced sections, \( bin_{1,2,\ldots,m} \), of width \( W_j \) contains at least \( k_{max} \) occurrences is equal to the union of the probability that each section contains at least \( k_{max} \) occurrences shown by the equation:

\[
P_j(\text{Any bin } \geq k_{max}) = P_j(bin_1 \geq k_{max} \cup bin_2 \geq k_{max} \cup \ldots \cup bin_m \geq k_{max}); \quad \text{eq}(8).
\]

Using the inclusion-exclusion principle of probability, we can decompose Equation 8 as follows:

\[
P_j(bin_1 \geq k_{max} \cup bin_2 \geq k_{max} \cup \ldots \cup bin_m \geq k_{max}) = \cdots
\]

\[
\cdots = P(bin_1 \geq k) + P(bin_2 \geq k) + \cdots + P(bin_m \geq k) \cdots
\]

\[
\cdots - P(bin_1 \geq k \cap bin_2 \geq k) - P(bin_1 \geq k \cap bin_3 \geq k) - \cdots
\]

\[
\cdots - P(bin_2 \geq k \cap bin_3 \geq k) \cdots - P(bin_{m-1} \geq k \cap bin_m \geq k) \cdots
\]

\[
\cdots + P(bin_1 \geq k \cap bin_2 \geq k \cap bin_3 \geq k) \cdots \text{ etc} \quad \text{eq}(9).
\]
Because all sections, \( bin_{1,2,\ldots,m} \), are of equal width \( W_j \) and the probability of incrementing each section is equal under the null hypothesis (McFadden, 1962), we can simplify Equation 8 as follows:

\[
P_j(bin_1 \geq k_{max} \cup bin_2 \geq k_{max} \cup \ldots \cup bin_m \geq k_{max}) = \ldots
\]

\[
= \sum_{k=k_{max}}^{n} \left\{ \begin{array}{l}
(-1)^2 P(1 \text{ bin} = k) \cdot \text{Number of 1 Bin Combinations} \\
\ldots + (-1)^3 P(2 \text{ bins} = k) \cdot \text{Number of 2 Bin Combinations} \\
\ldots + (-1)^4 P(3 \text{ bins} = k) \cdot \text{Number of 3 Bin Combinations} \\
\ldots + (-1)^{m+1} P(m \text{ bins} = k) \cdot \text{Number of m Bin Combinations}
\end{array} \right\} \quad \text{eq}(10)
\]

that can be further reduced to:

\[
P_j(\text{Any bin} \geq k_{max}) = \sum_{k=k_{max}}^{n} \left\{ \sum_{i=1}^{m} (-1)^{i+1} \binom{m}{i} P(i \text{ bins} = k) \right\} \quad \text{eq}(11).
\]

where \( n \) is the number of firing instances from the reference MUAPT, or total number of possible occurrences within a section, or \( bin \), of width \( W_j \). The probability that \( i \) bins have \( k \) occurrences can be found using Bayes conditional probability rule as follows:

\[
P(i \text{ bins} = k) = P(i - 1 \text{ bins} = k) \cdot P(bin_i = k | i - 1 \text{ bins} = k)
\]

\[
= P(bin_1 = k) \cdot P(bin_2 = k | 1 \text{ bin} = k) \cdot \ldots
\]

\[
\ldots \cdot P(bin_3 = k | 2 \text{ bins} = k) \cdot P(bin_i = k | i - 1 \text{ bins} = k) \quad \text{eq}(12)
\]

where:

\[
P(bin_1 = k) = \binom{n}{k} p_1^k (1 - p_1)^{n-k} \quad \text{eq}(13)
\]

and:

\[
P(bin_i = k | i - 1 \text{ bins} = k) = \binom{n - (i - 1)k}{k} p_i^k (1 - p_i)^{n-ik} \quad \text{eq}(14)
\]

where the probability \( p_i \) is defined by:

\[
p_i = \frac{1}{[m - (i - 1)]} \quad \text{eq}(15)
\]
In such a manner we calculated the probability $P_j$ of finding at least $k_{max}$ occurrences within any window of width $W_j$.

We repeated the above procedure using windows with progressively greater widths defined by equation 7. For each window of width $W_j$, we found the latency $L$ that provided the maximum number of occurrences $k_{max}$. Subsequently we computed the probability $P_j$ of finding $k_{max}$ occurrences within each window of width $W_j$. After repeating the procedure for all possible window widths, we identified the $j^{th}$ window that provided the minimum probability $P_{min}$ from all measured probabilities $P_j$. If the minimum probability, $P_{min}$, was less than the 0.05 significance threshold, then that pair of MUAPTs was determined to have statistically significant synchronization within the given interval of recurrence times tested.

Notice that Equations 12, 13 and 14 are well suited for detecting a statistically significant number of synchronized occurrences at relatively high significance levels, such as the 0.05 level used in this study. These equations approach the boundary of their efficacy when evaluating the statistical significance for a measured number of recurrence times relatively close to the number of recurrence times expected by chance (see Equation 3). However, under these conditions, the similarity between the number of recurrence times measured and the number of recurrence times expected by chance would indicate that the firing instances from the two motor units do not occur with statistically significant dependence. Therefore, this boundary condition has virtually no effect on the detections of statistically significant synchronization reported in this study.
REFERENCES


34. Keen DA and Fuglevand AJ. Common input to motor neurons innervating the same and
different compartments of the human extensor digitorum muscle. *Journal of

35. Keen DA, Chou LW, Nordstrom MA and Fuglevand AJ. Short-term synchrony in
diverse motor nuclei presumed to receive different extents of direct cortical input.

36. Kirkwood PA and Sears TA. The synaptic connexions to intercostal motoneurones as
revealed by the average common excitation potential. *Journal of Physiology* 275: 103-

37. Kirkwood PA, Sears TA, Tuck DL and Westgaard RH. Variations in the time course
of the synchronization of intercostal motoneurons in the cat. *Journal of Physiology* 327:
105-135, 1982.

38. Kline JC and De Luca CJ. Error Reduction in EMG Signal Decomposition. *(In Review)*

39. Kwiatkowski D, Phillips PCB, Schmidt P and Shin Y. Testing the null hypothesis of
stationarity against the alternative of a unit root. *Journal of Econometrics* 54: 159-178,
40. **Lippold OCJ, Redfearn JWT and Vuco J.** The electromyography of fatigue. 


FIGURE CAPTIONS

Figure 1. Recurrence times measured between several firing instances from one pair of stationary motor unit action potential trains (MUAPTs). Each MUAPT was designated as either a reference or an alternate MUAPT, with the reference MUAPT having fewer firing instances. We measured recurrence times between each firing instance in the reference MUAPT and the forwards and backwards firing instances of the alternate MUAPT, denoted by \( t_f \) and \( t_b \) respectively. The first order recurrence times are indicated by \( t_{f1} \) and \( t_{b1} \). Similarly the second and third order recurrence times are given by \( t_{f2} \) and \( t_{b2} \) and by \( t_{f3} \) and \( t_{b3} \), respectively. In total, five orders of recurrence times were measured for each firing instance in the reference MUAPT.

Figure 2. A depiction of our test for synchronization between pairs of stationary motor unit action potential trains (MUAPTs) with dependent firing instances using the SigMax detection method. A) The empirical cumulative distribution function of recurrence times (solid line) from one pair of MUAPTs is superimposed over the uniform cumulative distribution function (dashed line) predicted for MUAPTs with independent firing instances. A goodness-of-fit test was performed between the empirical and uniform cumulative distributions following the procedure detailed in step 2 of the Appendix. Deviation of the empirical cumulative distribution from the predicted uniform cumulative distribution beyond the 0.05 significance level calculated by the Cramer VonMises test statistic indicated that the two stationary MUAPTs were correlated and produced dependent firing instances. B) An example of the most statistically significant density of recurrence times detected from the empirical cumulative distribution function (solid line) beyond the 0.05 significance level using the equations provided in step 3 the Appendix. The
synchronization peak width $W$, shown as the width of the shaded box, and latency $L$, located at the center of the peak width, were recorded for further analysis. The total number of synchronized occurrences, $k_{\text{max}}$, was measured from the amplitude gained by the empirical cumulative distribution function within the shaded region. The number of synchronized occurrences expected due to chance, $\bar{k}$, was measured from the total amplitude gained by the expected uniform cumulative distribution function (dashed line) within the shaded region. Both $k_{\text{max}}$ and $\bar{k}$ were used to compute the synchronization index (SI) as the proportion of synchronized firing instances that occurred beyond what would be expected by chance (Equation 2).

Figure 3. Inter-pulse interval (IPI) histograms from 6 example MUAPTs: A) one from subject S1 contracting the FDI at 5% MVC; B) one from subject S3 contracting the FDI at 30% MVC; C) one from subject S4 contracting the FDI at 20% MVC; D) one from subject S5 contracting the VL at 35% MVC; E) one from subject S6 contracting the VL at 40% MVC; and F) one from subject S2 contracting the VL at 20% MVC. The mean IPI is indicated by the white vertical dashed lines while the IPI standard deviation is provided by white vertical dotted lines for each histogram. The D’Agostino-Pearson omnibus test revealed that the IPIs from all 6 histograms significantly deviated from a normal distribution ($p<0.0001$). The data are positively skewed (skewness>0) with relatively greater occurrences in the distribution tails (kurtosis>3) than would be expected by normal statistics.

Figure 4. Summary of synchronization (Sync.) measured by: the SigMax detection method (bars with no fill); the z-score synchronization detection method (bars with diagonal
The cumulative sum (cusum) synchronization detection method (bars with dotted fill); and the common input assumption synchronization detection method (bars with diamonds). Displayed is A,B) the percent of stationary motor unit action potential train (MUAPT) pairs with synchronization, C,D) the synchronization index (SI), E,F) the synchronization peak width (PW) and G,H) the synchronization latency (LAT) computed by each detection method for A,C,E,G) FDI and B,D,F,H) VL data. The bars marked with an asterisk indicate values that were not measured but assumed as prescribed by previously reported methods. For plots C-H) the bars indicate the 95% confidence interval above and below the average value of each synchronization metric shown by the horizontal line.

**Figure 5.** Differences in synchronization (Sync.) measured between the same set of MUAPTs obtained with and without error reduction. Synchronization was evaluated using the SigMax detection method from pairs of A) FDI and B) VL MUAPTs. Displayed is the average magnitude of the percent difference (Diff.) of additional (positive) and missed (negative) synchronization detections. Also plotted are the average percent difference (horizontal line) and 95% confidence interval (bars) of the difference in the synchronization index (SI), the synchronization peak width (PW) and the synchronization latency (LAT). These discrepancies represent the changes in the amount of synchronization that result from unmitigated decomposition errors.

**Figure 6.** Synchronization (Sync.) measured from the cross-correlation histogram of recurrence times using a z-score significance threshold. The mean bin amplitude is shown by the horizontal dashed line and the normal significance threshold for a z-score >1.96 is shown by the
horizontal solid line. In total, 12 separate peaks exceeded the normal significance threshold indicating 12 separate detections of synchronization. This compares to the single statistically significant peak detected by our SigMax detection method, illustrated by the region of the histogram shaded gray.

Figure 7. Synchronization (Sync.) discrepancies evaluated between the SigMax detection method and one of three other synchronization methods tested: the z-score synchronization detection method (bars with diagonal lines); the cumulative sum (cusum) synchronization detection method (bars with dotted fill); and the common input assumption synchronization detection method (bars with diamonds). Differences were computed between the same set of stationary A,C,E,G) FDI and B,D,F,H) VL MUAPTs. Displayed is A,B) the average magnitude of the percent difference (Diff.) of additional (positive) and missed (negative) synchronization detections. Also shown are the average percent difference (horizontal line) and 95% confidence interval (bars) of the difference in C,D) the synchronization index (SI), E,F) the synchronization peak width (PW) and G,H) the synchronization latency (LAT).

Figure 8. Examples of synchronization measured using the cumulative sum detection method from two different pairs of stationary motor unit action potential trains (MUAPTs). For each pair, data in A,D) the cross-correlation histogram was used to compute B,E) the normalized cumulative sum of bin amplitudes. The gray dashed lines superimposed on the cross-correlation histograms indicate the mean values used to compute the cumulative sum. The 0.1 and 0.9 significance threshold of the cumulative sum data are shown by horizontal dashed lines in black. These thresholds were used to detect synchronization peaks. B) For MUAPT pair #1, no peak
was found within the first recurrence interval of the cumulative sum data. C) However, our SigMax detection method found a synchronization peak from the empirical cumulative distribution of recurrence times, indicated by the region shaded gray. E) For MUAPT pair #2, a 13 ms wide peak was detected between -9 ms and 4 ms from the cumulative sum data. F) Our SigMax detection method found a 22.7 ms wide peak from the empirical cumulative distribution of recurrence times, shown by the region shaded gray.

Figure 9. Examples of synchronization measured from the common input assumption synchronization detection method for two different pairs of stationary MUAPTs. A) For one pair the common input assumption detection method located the synchronization peak at 0 ms latency with a width of 11 ms, yielding a synchronization index of -8.14, indicated by the boxed region shaded by diagonal lines. However, our SigMax detection method found no statistically significant synchronization peak for the same data. B) For another MUAPT pair the common input assumption synchronization detection method assumed a synchronization peak at 0 ms latency with a width of 11 ms, yielding a synchronization index of 0.536, indicated by the boxed region shaded by diagonal lines. Our SigMax detection method found a synchronization peak located at 10.4 ms latency with a peak width of 6.78 ms and synchronization index of 5.88, indicated by the region shaded in gray.

Figure 10. Distributions of the percentage of stationary motor unit action potential train (MUAPT) pairs with synchronized firing instances within each recurrence interval for A) FDI and C) VL data. The total amplitude of each bar indicates the total percentage of paired MUAPTs with synchronization. The amplitude of solid gray bars shows the percentage of
MUAPT pairs with synchronization that did not result from harmonics (see explanation in the results). From these data, all recurrence intervals outside of the first interval contained fewer than 4% of pairs of MUAPTs with synchronization, indicating a non-statistically significant quantity of synchronization detections. For the remaining statistically significant synchronization detections the distributions of the synchronization latency are shown for B) FDI and D) VL pairs of stationary MUAPTs. Overall, the synchronization latency data ranged from -7.4 to 7.2 and from -6.0 to 5.5 ms for 95% of FDI and VL data respectively.

Figure 11. The inter-pulse intervals (IPIs) of two example MUAPTs measured during the 25 second constant force region of the contraction. Shown are A) the IPIs of a stationary MUAPT and B) the IPIs of a moderately non-stationary MUAPT. The KPSS test statistic indicated in the upper right corner of each plot was used to determine the stationarity of each MUAPT. A KPSS test statistic greater than 0.463 indicated the MUAPT contained statistically significant non-stationary firing instances.

Figure 12. A graphical depiction of the third step of the SigMax detection method. Synchronization was measured from the empirical cumulative distribution of recurrence times between pairs of stationary MUAPTs with dependent firing instances. A window of width $W_j$ was selected and used to calculate the number of recurrence times at different latencies, $L$. The latency that provided the maximum number of recurrence times, $k_{max}$, was selected for significance testing. Statistical significance was computed as the probability, $P_j$, of finding $k_{max}$ occurrences within a window of width $W_j$, using the equations in step 3 of the Appendix. This procedure was repeated for windows of different widths defined by Equation 7 in the text.
The most statistically significant density of recurrence times was identified as the $j^\text{th}$ window width that produced the minimum probability, $P_{min}$, of all computed probabilities, $P_j$. If $P_{min}$ was less than the 0.05 significance threshold then the pair of MUAPTs manifested statistically significant synchronization.
**Table 1. Statistical results of the normality test for motor unit IPIs**

<table>
<thead>
<tr>
<th>Muscle</th>
<th>Non-normal MUAPT IPIs (n)</th>
<th>Significance</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDI</td>
<td>98.6% (794)</td>
<td>0.00255</td>
<td>0.46</td>
<td>6.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0, 0.0294]</td>
<td>[-0.63, 2.02]</td>
<td>[2.95, 17.6]</td>
</tr>
<tr>
<td>VL</td>
<td>99.8% (1,206)</td>
<td>0.00026</td>
<td>0.45</td>
<td>6.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0, 0.00897]</td>
<td>[-0.97, 2.08]</td>
<td>[3.21, 17.2]</td>
</tr>
</tbody>
</table>

*Table 1. Results from the D’Agostino-Pearson omnibus test for normally distributed IPIs from stationary FDI and VL motor unit action potential trains (MUAPTs). The percentage of MUAPTs with IPIs that significantly deviated from a normal distribution is shown as well as the mean and 95% confidence interval of the D’Agostino-Pearson omnibus test significance, skewness and kurtosis. Significance values below 0.05 indicated non-normal data; skewness values greater than 0 indicated positive skewness; kurtosis values greater than 3 indicated relatively greater occurrences in the tails of the distribution than would be expected by normal statistics.*
Figure 1

Measuring Recurrence Times

Time (s)

Reference MUAPT
Alternate MUAPT

$\text{Measuring Recurrence Times}$

$t_{b1}$ $t_{f1}$

$t_{b2}$ $t_{f2}$

$t_{b3}$ $t_{f3}$
SigMax Synchronization Detection

Evaluating Dependence Between Stationary MUAPTs

A)

Cramer VonMises Goodness-of-fit

Statistically Significant Synchronization Detection

B)

Recurrence Interval
Latency (ms)

\[ i \times \text{TP}_{\text{Alt}} \pm \frac{\text{TP}_{\text{Alt}}}{2} \]

\[ i^{th} \text{ Recurrence Interval Latency (ms)} \]
Figure 3

Motor Unit Firing Instances Have Non-normal Distribution

A) S1 FDI 5% MVC MU1
   Skewness = 0.60
   Kurtosis = 5.64
   p<0.00001

B) S3 FDI 30% MVC MU24
   Skewness = 0.50
   Kurtosis = 6.26
   p<0.00001

C) S4 FDI 20% MVC MU25
   Skewness = 0.29
   Kurtosis = 4.86
   p<0.00001

D) S5 VL 35% MVC MU11
   Skewness = 0.19
   Kurtosis = 4.20
   p<0.0001

E) S6 VL 40% MVC MU17
   Skewness = 0.48
   Kurtosis = 5.27
   p<0.00001

F) S2 VL 20% MVC MU32
   Skewness = 0.98
   Kurtosis = 4.56
   p<0.00001
Figure 4

Summary of Synchronization Measured from 4 Different Detection Methods

- FDI Sync. MUAPTs
- VL Sync. MUAPTs
- FDI Sync. Index
- VL Sync. Index
- FDI Peak Width
- VL Peak Width
- FDI Latency
- VL Latency

<table>
<thead>
<tr>
<th>Detection Method</th>
<th>SigMax</th>
<th>Z-Score</th>
<th>Cusum</th>
<th>Common Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDI Sync. MUAPTs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VL Sync. MUAPTs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDI Sync. Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VL Sync. Index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDI Peak Width</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VL Peak Width</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDI Latency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VL Latency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5

Synchronization Discrepancies
Resulting from Decomposition Errors

A) FDI Sync. Discrepancies

B) VL Sync. Discrepancies

Sync Detections

SI  PW  LAT

Difference (%)
Figure 6

Z-Score Synchronization Detection Method

- Z-Score Detection
- SigMax Detection
Figure 7

Synchronization Discrepancies from 3 Different Detection Methods Relative to SigMax

FDI Sync. Detection Discrepancies

FDI Sync. Index (SI) Discrepancies

FDI Peak Width (PW) Discrepancies

FDI Latency (LAT) Discrepancies

VL Sync. Detection Discrepancies

VL Sync. Index (SI) Discrepancies

VL Peak Width (PW) Discrepancies

VL Latency (LAT) Discrepancies

Z-Score  Cusum  Common Input

Z-Score  Cusum  Common Input
Figure 8

Cumulative Sum Synchronization Detection From MUAPT Pair #1

Cross-Correlation Histogram

Cumulative Sum

SigMax

Cumulative Sum Synchronization Detection From MUAPT Pair #2

Cross-Correlation Histogram

Cumulative Sum

SigMax
Figure 9

**Common Input Assumption Detection of Negative Synchronization**

- SI = -8.14

**Common Input Assumption Peak Location Difference**

- SI = 0.54
- SI = 5.88
Figure 10

Statistical Factors Responsible for Long-Term Synchronization

Sync. Harmonics

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>No Harmonics</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) FDI</td>
<td><img src="image" alt="Graph A) FDI" /></td>
<td></td>
</tr>
<tr>
<td>C) VL</td>
<td><img src="image" alt="Graph C) VL" /></td>
<td></td>
</tr>
</tbody>
</table>

Sync. Harmonics Removed

<table>
<thead>
<tr>
<th></th>
<th>Number of MUAPT Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>B) FDI</td>
<td><img src="image" alt="Graph B) FDI" /></td>
</tr>
<tr>
<td>D) VL</td>
<td><img src="image" alt="Graph D) VL" /></td>
</tr>
</tbody>
</table>

Recurrence Interval of Alt MU mean IPI

-5 -4 -3 -2 1 2 3 4 5

Number of MUAPT Pairs

-20 -10 0 10 20

Peak Latency (ms)
Figure 11

Testing Motor Units for Stationary Firings

Stationary MU IPIs

Non-stationary MU IPIs

KPSS = 0.06

KPSS = 1.19
Figure 12

Finding $k_{max}$

$W_j$ \\

$k_{max}$ \\

$L_1 - L_2 - L_3$ \\

Probability \\

Occurrences \\

$i \times TP_{alt} \pm \frac{TP_{alt}}{2}$ \\

$i^{th}$ Recurrence Interval \\
Latency (ms)