Bayesian Network Modeling for Discovering "Dependent Synergies" among Muscles in Reaching Movements

Junning Li[1], Z. Jane Wang[1], Janice J. Eng[2] and Martin J. McKeown[3,4]

[1] Department of Electrical and Computer Engineering, [2] School of Rehabilitation Sciences,
[3] Department of Medicine (Neurology) and [4] Pacific Parkinson's Research Centre,
University of British Columbia, Vancouver, Canada

***Abstract*—The coordinated activities of muscles during reaching movements can be characterized by appropriate analysis of simultaneously-recorded surface Electromyograms (sEMGs). Many recent sEMG studies have analyzed muscle synergies using statistical methods such as Independent Component Analysis, which commonly assume a small set of influences upstream of the muscles (e.g., originating from the motor cortex) produce the sEMG signals. Traditionally only the amplitude of the sEMG signal was investigated. Here we present a fundamentally different approach and model sEMG signals after the effects of amplitude have been minimized. We develop the framework of Bayesian networks (BNs) for modelling muscle activities and for analyzing the overall muscle network structure. Instead of assuming that synergies may be independently activated, we assume that neuronal activity driving a given muscle may be conditionally dependent upon neurons driving other muscles. We call the resulting interactions between muscle activity patterns "dependent synergies". The learned BN networks were explored for the purpose of classification across subjects based on hand dominance or affliction by stroke. Network structure features were investigated as classification input features and it was determined that specific edge connection patterns of 3-node subnetworks were selectively recruited during reaching movements and were differentially recruited after stroke compared to normal control subjects. The resulting classification was robust to inter-subject and within-group variability and yielded excellent classification performance. The proposed framework extends muscle synergy analysis and provides a framework for thinking about muscle activity interactions in motor control.**

***Index Terms*—surface Electromyogram (sEMG), Bayesian network, reaching movement, synergy, network analysis**

I. Introduction

An important goal of motor control studies is to understand how the central nervous system (CNS) selects and co-ordinates the muscle activity patterns necessary to achieve a variety of natural motor behaviors [1]. A key emerging concept in motor control is the importance of synergies [1], or groups of muscles that act together. Works on frogs have suggested that a complex repertoire of movements can emerge from the appropriate control and selection of only a few synergies which each represent a primitive movement [2]. However, identifying muscle synergies from all possible muscle patterns and efficient decomposing of complex and variable motor behaviors into meaningful synergies remain challenging problems. To address this goal, a necessary intermediate step is to determine how muscles efficiently collaborate together during movements. In this paper, we plan to infer muscle interaction patterns from surface Electromyogram (sEMG) recordings during reaching movements. Especially we are interested in investigating whether certain muscle interactions are selectively recruited across subjects based on hand dominance or affliction by stroke.

A sEMG is a semi-stochastic signal whose properties depend upon a number of factors including the anatomical and physiological properties of the contracting muscles, the amount of subcutaneous fat, and choice of electrodes [3]. Standard analytical methods, including frequency-based ones, may be particularly sensitive to parameters difficult to measure, such as capacitive effects of muscles and subcutaneous tissue [4]. Nevertheless, despite its limitations, the non-invasive nature of sEMG makes it practical to record several muscles simultaneously in humans, and hence allows the investigation of synergies.

A sEMG signal can be modelled as a zero-mean wide-sense stationary stochastic process (*i.e.* the so-called carrier signal) modulated by the sEMG amplitude [5]. A common practice in the sEMG literature is to focus only on the amplitude data, *i.e.* doing rectifying and then low-pass filtering the raw sEMG signal, while the carrier data is generally ignored. Here we use a fundamentally different approach and model sEMG carrier signal after the effects of amplitude have been minimized.

During the last years, partially linear decomposition methods such as Principal Component Analysis (PCA), linear Independent Component Analysis (ICA) [6] [7], and Non-negative Matrix Factorization (NMF) [8] have been suggested to infer synergistic action between muscles. For instance, linear ICA was applied to noisy sEMG data and revealed meaningful interactions between muscles [6]. These methods are characterized by a number of latent variables and project multi-channel sEMG signals to a subspace. They share a common assumption that a small set of source signals are upstream of the muscles and produce the sEMG signals. For example, ICA assumes that the observations are combinations of statistically independent components, and the goal of ICA

Contact information: Z. Jane Wang, UBC, 2332 Main Mall, Vancouver, BC V6T 1Z4, Canada. Email: zjanew@ece.ubc.ca.

is to find the underlying sources.

Here we propose a different approach, the Bayesian network (BN) modeling approach, to directly represent interactions between muscles without using latent variables. Methods such as P/ICA do not explicitly reveal interactions but implicitly through underlying sources. Our BN framework provides an alternative which captures the interactions directly from the observed sEMG signals by detecting conditional dependence/independence between muscle activities, *i.e.* whether the activities of two muscles are associated given that of a third muscle.

There are a number of biological and statistical reasons that make the assumption of conditional dependence between muscle activities plausible. Depending upon the neural context, the same neurons participating in central pattern generators can demonstrate remarkably dissimilar behaviors [9]. Activity in muscles themselves may be modulated by Ia inhibitory interneurons from antagonists or Renshaw cell activity in the spinal cord that may project to other motor neurons in the spinal cord. [10]. Ordinary coherence, often used to infer connectivity coupling between muscles, cannot distinguish whether two channels are directly connected or indirectly connected via other channels suggesting that conditional dependence may need to be considered [11]. In fact, failure to identify potential conditional dependence between muscles may lead to erroneous interpretations regarding the overall interactions between muscles, necessitating the use of partial coherence, often resulting in a distinctly different connection pattern [11]. A BN represents the conditional dependence/independence through a graph of nodes and edges connected according to rigorous statistical rules (see Sec. II-B), so it is suitable to discover conditional interactions between muscles.

There are a number of reasons why BNs, as opposed to other choices of graphical models (such as Boolean networks), may be particularly well-suited for modeling muscle interaction networks. First, BN models have a solid basis in statistics, enabling them to deal with the stochastic and nonlinear aspects of sEMG measurements in a natural way. As a rigorous probabilistic model, a BN allows incorporation of the stochastic nature of sEMG recordings that may be caused by any number of biological factors along the cortex $\rightarrow$ spinal cord $\rightarrow$ peripheral nerve $\rightarrow$ neuromuscular junction $\rightarrow$ muscle pathway. Second, BN's modular nature makes it easily extensible to the task of modeling sub-networks of sEMG signals. In BNs, conditional probability distributions (CPD) are specified locally at each node to encode dependence relationships between a node and its parents, so the whole network can be decomposed into many small sub-networks. Since a node is independent of its ancestors given its parents, a simple conditional independence relationship between muscles could be that the interaction between muscle-A and muscle-B does not depends on the activity in muscle-C. Further, the rich repertoire of techniques developed for network analysis in other areas can be used for inferring muscle networks using sEMG. Finally, BNs can be used when incomplete knowledge is available, and can also deal with dynamical aspects of muscle interactions through generalizations like dynamical BNs.

Therefore, in this paper, we develop a BN modeling framework to statistically capture the interactions between muscle activity patterns directly. More specifically, we generalize the muscle synergy idea into the concept of a *muscle network*, defined as a set of muscle activity patterns with *probabilistic and conditional* interactions between them that are coordinated to achieve specific motor behaviors. In our approach, we first model the overall muscle activity across several simultaneously recorded sEMG signals. From the learned muscle network we then define muscle synergies as statistically significant sub-networks or network-motifs [12] and use the term "dependent synergies" to refer to (conditionally) dependent muscle (not necessarily pair-wise) interactions.

We also plan to tackle the problem of investigating consistent muscle synergies across subjects within a certain group. While there are many potential factors that may affect muscle synergies and thus sEMG patterns during a reaching movement, here we focus on the effects of stroke and hand dominance. Hand dominance has been reported as an important factor in motor control [13], and different muscle activations between the dominant and non-dominant hands have recently been observed during reaching movements [14]. Though hand dominance as a factor in motor and functional performance has been studied in the literature, to the best of our knowledge, no studies have investigated the impact of hand dominance in healthy or stroke subjects in terms of muscle association/interaction patterns.

The main contributions of this paper are as follows:

- To present a framework for learning the *muscle interaction networks* during reaching movements based on the BN modeling of sEMG data.
- To demonstrate how the trained BNs can then be probed with network motif analysis to determine "dependent synergies".
- To demonstrate that some network structure features are relatively robust across subjects, and thus can be used to distinguish factors such as handedness and stroke status.
- To demonstrate that specific three-muscle synergies may provide insights into the compensatory changes seen in reaching movements after stroke.
- To indicate that the sEMG "carrier" signal (after the amplitude information is estimated and removed) can also be informative, even though it has been traditionally ignored by sEMG analysis methods.

The paper is organized as follows. In Section II, we describe the proposed BN framework for learning muscle networks and analyzing the sub-network patterns. A real case study utilizing sEMG recordings from stroke and healthy subjects, including data from both dominant and non-dominant arms, is discussed in Section III. Finally, we conclude our paper and suggest some directions for future research.

II. METHODS

A. Framework and its components

Our Bayesian network (BN) framework includes three components: BN modelling, graph structure analysis and classification. First, BNs are applied to multi-muscle sEMG signals,

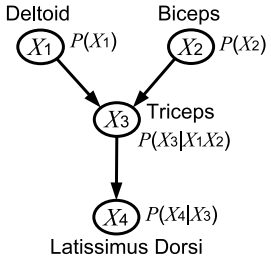


Fig. 1. An example of representing the stochastic interactions of sEMG recordings of different muscles with a Bayesian Network. This particular graph is only for illustrative purposes and does not have intrinsic biological meaning. Each node represents a sEMG recording of a muscle and the whole graph represents the interaction among the muscles in a movement. This DAG encodes the independence relationships $X_1 \perp X_2$ (but $X_1 \not\perp X_2|X_3$) and $(X_1, X_2) \perp X_4|X_3$ (but $(X_1, X_2) \not\perp X_4$). Nodes X_1 and X_2 do not have any parent and they are associated with unconditional probability distributions. Nodes X_3 and X_4 have parent(s) and they are associated with conditional probability distributions. The joint probability distribution can be factorized according to the DAG as: $P(X) = P(X_1)P(X_2)P(X_3|X_1X_2)P(X_4|X_3)$.

with directed acyclic graphs (DAG) encoding the overall interactions between muscles. Researchers can choose different types of BNs according to their interest and prior knowledge about a specific application. For example, they can use a static BN to model the invariant interactions, a dynamic BN [15] to model the dynamics of muscles, or a BN with hidden nodes [15] to model the unobserved neural signals which drive the muscles.

Secondly, graph structure analysis is conducted on the learned BNs to extract structural features which characterize the interaction patterns among muscle activity patterns. Structural features can be the number of edges in or out from a node (which is called "degree" in graph theory), or the length of the shortest path from one node to another (which is called "distance" in graph theory). To go beyond the node and edge levels, but to a sub-network level, we employed the "network motif" concept [12].

Thirdly, classification is performed, based on the BNs. As statistical models, BNs can be naturally extended to statistical classifiers with the posterior probability criterion. Alternatively, the DAGs of BNs can also be used as input features to other classifiers such as classification trees. In the following sub-sections, we will elaborate on each of the three components alluded to above.

B. Bayesian networks

a) Introduction: A Bayesian Network (BN) [16], also referred as a "Bayesian belief network" or simply "belief network", is a graphical model that consists of a *directed acyclic graph* (DAG) and a set of (conditional) probability distributions. The DAG encodes the (conditional) dependence/independence relationships among random variables, and the probability distributions constitute the joint probability distribution with Bayes' rule. A BN, in short, is a representation of the joint distribution over random variables by indicating the conditional dependence/independence relationships with a DAG.

A DAG encodes a set of (conditional) independence relationships between node variables with the concept of d-separation [16]. For instance, let X_1, X_2 and X_3 denote three node variables. According to the global Markov property [16], if X_1 is d-separated from X_2 by X_3 in the DAG, then it is said that X_1 is conditionally independent of X_2 given X_3, *i.e.* $P(X_1X_2|X_3) = P(X_1|X_3)P(X_2|X_3)$, and we denote it as $X_1 \perp X_2|X_3$. The definition of conditional independence is similar to that of unconditional independence $P(X_1X_2) = P(X_1)P(X_2)$ except that it is conditional on a third random variable X_3. Here we focus on two simplified but important corollaries within the broader concept of d-separation. (1) If X_1 and X_2 are connected (*i.e.* there is a path in the DAG from X_1 to X_2 or vice versa), then X_1 is not independent of X_2 which we denote as $X_1 \not\perp X_2$. (2) If X_1 precedes X_2 (*i.e.* there is at least a path from X_1 to X_2) and X_3 blocks all the paths from X_1 to X_2, then $X_1 \perp X_2|X_3$. Fig. 1 shows an example of BN which represents a network of four muscles. (Note that this figure is only for illustrative purposes and does not necessarily represent any real muscle network.) According to the above corollaries, this DAG encodes a set of relationships: $X_1 \perp X_2$ (but $X_1 \not\perp X_2|X_3$) and $(X_1, X_2) \perp X_4|X_3$ (but $(X_1, X_2) \not\perp X_4$). This example also shows that conditional independence does not imply unconditional independence, and vice versa.

If the joint probability distribution of a set of random variables X is subject to the Markov property, it can be factorized according to the DAG as

$$P(X) = \prod_{\text{pa}[X_i]\neq\emptyset} P(X_i|\text{pa}[X_i], \theta_i) \prod_{\text{pa}[X_i]=\emptyset} P(X_i|\theta_i), \quad (1)$$

where the set $\theta = \{\theta_1 \ldots \theta_n\}$ denotes the parameters used in the probability distributions and pa$[X_i]$ denotes parents of X_i, *i.e.* nodes with an edge to X_i. If a node X_i has parent nodes, *i.e.* pa$[X_i] \neq \phi$, it is associated with a conditional probability distribution $P(X_i|\text{pa}[X_i])$. If a node X_i does not have any parent, it is associated with an unconditional probability distribution $P(X_i)$. For example, the joint probability of the four-muscle network in Fig. 1 generally can be decomposed as Eq. 2 according to the chain rule, and further as Eq. 3 according to the (conditional) independence relationships specified by the DAG.

$$P(X) = \prod_{i=1}^{4} P(X_i|X_1, \ldots, X_{i-1}) \quad (2)$$

$$= P(X_1)P(X_2)P(X_3|X_1X_2)P(X_4|X_3). \quad (3)$$

Using BNs to represent muscle interactions can be summarized as follows. sEMG signals are regarded as a vector-valued stochastic process $X(t)$=[$X_1(t)$, $X_2(t)$, $\ldots X_n(t)$]^T where n denotes the number of muscles and $X_i(t)$ the observed signal of the ith muscle at time t. A DAG with nodes $X = \{X_1 \ldots X_n\}$ indicates the interactions between muscles. If a node X_i is connected to another node X_j, then the corresponding muscles are considered to interact. If X_i is d-separated from X_j by another node X_k, then the muscles represented by X_i and X_j do not interact conditionally on the activity of the muscle X_k. Conditional and unconditional probability distributions are associated with the nodes, describing how

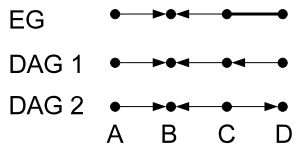


Fig. 2. An example of essential graphs (EGs). Both the directed and acyclic graphs DAG 1 and DAG 2 represent the same set of conditional independence: $A \perp (C, D)$, $B \perp D | C$ and $B \perp D | (A, C)$. Since the edge between C and D is reversible, its direction is removed in the EG. Directions in either the DAGs or the EG do not necessarily imply causality.

the muscle activity patterns interact with each other. In this study, we employed Gaussian BNs to model the multi-muscle sEMG signals, *i.e.* we modelled each variable X_i as the sum of a Gaussian noise and a linear combination of its parents pa$[X_i]$. Gaussian BNs are not only applicable but also one of the most popular BNs for modeling multi-channel continuous variables.

It should be pointed out that before DAGs are analyzed to reveal muscle interactions, it must be converted to an *essential graph* (EG) [17] which is also referred as a *completed acyclic partially directed graph* (CPDAG) in the literature. This is because there are often several different DAGs representing the same set of conditional independence relationships, while there is only one EG uniquely encoding the set of conditional independence relationships, as shown in Fig. 2. An EG has the same edges as a DAG does except that some edges are not directed. Algorithms to convert a DAG to an EG have been proposed previously [17].

A directed graph, such as those that contain directed edges of either a BN or an EG to encode conditional independence/dependence and independence/dependence, does not necessarily imply causality but rather association. Hence our model dose not conflict with the fact that the activities of multiple muscles are often coupled by kinematics and dynamics of the bones and joints. For further details on BNs, EGs and d-separation, the reader is referred to Lauritzen's book (1996) [16] and Andersson's paper (1997) [17].

b) Learning Bayesian networks: Learning a BN includes two steps: (1) structure learning and (2) parameter learning. Structure learning is to select an appropriate DAG among many candidate DAGs. Parameter learning is to estimate the parameters of the conditional and unconditional distributions given the DAG. In structure learning, we attempted to select the most probable DAG based on the observations according to the maximum *a posteriori* (MAP) criterion. Let X denote the observations and S the DAG, the best structure from the view of Bayesian statistics is:

$$\hat{S} = \arg\max_{S} p(S|X), \tag{4}$$

where according to Bayes' rule, we have

$$p(S|X) = \frac{p(X|S)p(S)}{p(X)}, \tag{5}$$

$$p(X|S) = \int p(X|\theta, S)p(\theta|S)d\theta, \tag{6}$$

where $p(S)$ is the prior probability of the structure S and $p(\theta|S)$ is the probability of the parameter θ given the structure S. The MAP criterion has the advantage of allowing users to incorporate their knowledge in the prior probability. As the denominator in Eq. (5) does not depend on S, only the numerator needs to be maximized. If $p(S)$ (the prior probability) is uniform over all the possible structures, only $p(X|S)$ in Eq. (6) (the conditional probability) needs to be maximized. The uniform assumption is reasonable in practice since we do not prefer any structure before we observe the data. However, alternates to a uniform prior distribution may be considered to enhance computationally efficiency [18].

The MAP approach can be implemented by selecting the structure with the largest Bayesian Information Criterion (BIC) score [19] which is defined as

$$\text{BIC}(S) = \sup_{\theta} \log P(X|S, \theta) - 0.5K \log N, \tag{7}$$

where N denotes the sample size of X and K denotes the number of free parameters in θ. In the comparison between two models S_1 and S_2, exp[BIC(S_1)-BIC(S_2)] asymptotically approximates the ratio of their posterior probability $P(S_1|X)/P(S_2|X)$ if the two models S_1 and S_2 have the same prior probability, *i.e.* $P(S_1)$=$P(S_2)$ and $p(\theta|S)$ is uniform [19]. (For rigorous proof, please refer to Schwarz's paper in 1978.) The large sample size in our sEMG study, 1000 time points (see Sec. III-A), should satisfy the condition of the asymptotical approximation. As shown in Eq. (7), the BIC consists of two terms: the maximum log likelihood term $\sup \log P(X|S)$ and the penalty term $-0.5K \log N$. The penalty term prevents "over fitting", *i.e.* choosing a structure which has too many edges compared with the data size. Because a structure with more edges tends to have larger likelihood, we will inevitably choose a fully connected DAG if only the maximum likelihood criterion is used. Therefore, a penalty term is needed in model selection. The BIC penalty term is proportional to the number of free parameters (K), and "punishes" structures with redundant edges.

After structure learning, we estimated the parameters of the conditional and unconditional distributions via the maximum likelihood criterion. Since the number of all the possible DAGs is super-exponential to the number of nodes, it is impractical to exhaustively search for the best DAG. To avoid local maxima found by greedy algorithms, we employed the Markov chain Monte Carlo algorithm (MCMC) [20] to learn the structure.

Our implementation of learning BNs was developed based on the software Bayes Net Toolbox (BNT) [21] for Matlab.

C. Sub-network Patterns and Muscle Synergies

Milo and Shen-Orr [12] reported that in the networks of the real world (for instance, gene regulation networks), certain connection patterns of sub-networks appear more frequently than would be expected from chance, and these patterns (named "network motifs" by the authors) can be used to characterize the networks. Since in the current case, sub-networks represent the co-activation between muscles, we adopt a similar idea to discover the muscle synergies from the DAGs of BNs. However, in contrast to the original proposal by Milo and Shen-Orr's, where the goal was to determine network motifs that appear more frequently in real graphs

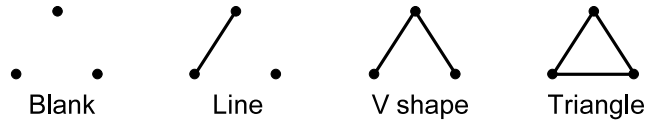


Fig. 3. Different ways that three muscles can interact within a BN framework. The directed triple graph is first converted to an undirected graph, and then it is classified as one of the patterns. The four patterns are abbreviated as B, L, V and T respectively.

than in randomized graphs, our goal is to identify network motifs that distinguish a group of graphs from another group. Specifically, the question of interest is that "given two groups of muscle interaction graphs derived from sEMG recordings (*e.g.* recordings of stroke and normal subjects), is it possible to determine which network motif(s) distinguish one group from the other?"

We propose the following way to detect network motifs distinguishing between two groups of graphs. First, the occurrences of each possible connection pattern in each graph is counted. As a result, the count of a particular connection pattern in one group of graphs is a group of numbers. Then, the two groups of numbers are compared with a hypothesis test such as a t-test. Finally, patterns appearing significantly more frequently in one group than in the other are selected as the feature network motifs of that group. We note that patterns which are functionally important but not statistically significant could exist and could be missed by this approach.

In our sEMG study, we focus on triplet network motifs, *i.e.* sub-networks with three nodes. Though sub-networks with more nodes can be analyzed similarly without theoretic difficulty, we did not pursue more than three in this exploratory research due to limited computation power and the observation that triplets have demonstrated our framework adequately. The complete and detailed procedure of detecting and evaluating triplets from two groups of graphs is as follows. First, DAGs are converted to EGs because a EG uniquely determines the dependence relationships among nodes (see Sec. II-B). Secondly, the appearing frequencies of triplet connection patterns are counted. The possible triplet patterns are show in Fig. 3. Thirdly, t-test is performed to evaluate whether a specific connection pattern appears significantly more often in one group of EGs than in another. As a result, each triplet pattern is associated with a level of significance, or p-value. Forth, the p-values are adjusted for the effect of multiple testings with Sidak correction as in Eq. (8),

$$p_a = 1 - (1-p)^h, \tag{8}$$

where p is the original p-value, p_a is the adjusted one and h is the number of hypotheses tested simultaneously. In the context, h equals the number of the interested connection patterns. The effect of multiple testings can also be adjusted with the false discovery rate (FDR) [22] which controls q-values, *i.e.* the expected portion of falsely rejected hypotheses among those rejected. Finally, connection patterns with the adjusted p-values or q-values lower than 0.05 are selected as network motifs.

While the above general network motifs provide information on the overall connectivity patterns of the network, we are also interested in specific muscle triplets because it is possible that alterations between the interactions of a few particular muscles may significantly influence classification of reaching movements between groups. The identification procedure is similar to that of the general network motifs, except that the connection patterns are counted for each combination of three specific muscles individually. For n muscles, all the C_n^3 triplets are exhaustively examined. Instead of a t-test, Fisher's exact test is employed to check whether a connection pattern of a specific muscle triple appears significantly more in a group of graphs than in another. The effect of multiple testing is also adjusted, but the number of simultaneous hypothesis tests h is much larger. If there are n muscles and m patterns of interest, the number h is mC_n^3.

D. Classification

BNs can be used for classification purpose in two ways: (1) as statistical models, they can be extended to be statistical classifiers; (2) as graphical models, their structures can be input as features to other classifiers.

BNs can be extended to be a statistical classifier naturally with the posterior probability criterion. Suppose M_1 and M_2 are the statistical models of the sEMG signals of two groups of subjects respectively, *e.g.* a control group and a stroke group. Given sEMG signals X, its model index can be predicted by the posterior probability criterion as in Eq. (9). The more $p(M_1|X)$ is larger than $p(M_2|X)$, the more likely that X belongs to group 1, and vice versa. If the prior probabilities $p(M_1)$ and $p(M_2)$ are equal, the ratio $p(M_1|X)/p(M_2|X)$ is the same as the ratio $p(X|M_1)/p(X|M_2)$ according to Bayes' rule. Suppose there are totally N_i subjects of group i and M_{ij} represents the BN model of the jth subject in group i. With the assumption that each individual model of group i is equally representative of the group, the group model M_i (i=1,2) can be built by averaging the BN models trained from individual subjects within the same group, as expressed in Eq. (10).

$$\text{If } \frac{p(M_1|X)}{p(M_2|X)} \begin{cases} > 1, \text{ then X belongs to group 1,} \\ < 1, \text{ then X belongs to group 2;} \end{cases} \tag{9}$$

$$p(X|M_i) = \frac{\sum_{j=1}^{N_i} p(X|M_{ij})}{N_i}, i = 1, 2. \tag{10}$$

Structure features of BNs can also be used as the input to other classifiers in various ways. As we mentioned in Sec. II-B, DAGs should be converted to EGs before used to represent the interactions between muscles. An EG of n nodes can be encoded as an n-by-n binary adjacent matrix $A = \{a_{ij}\}$ where $a_{ij} = 1$ indicates an edge from node i to node j. Because an EG cannot have any edge which circles from and to the same node, the diagonal elements are all zeros, and they are uninformative. The elements off the diagonal line can be lined up as a binary vector with $n(n-1)$ elements and then used as input to a classification tree. We choose a classification tree, but not other classifiers such as the support vector machine (SVM) [23] because of two reasons. First, classification tree is especially suitable for categorical data as an adjacent matrix is; secondly and also

most importantly, it is easier to interpret since a classification tree explicitly gives the conditions of predicting the class. In contrast, despite its popularity, SVM results are hard to interpret, as the SVM algorithm implicitly maps features to a high-dimensional imaginary space. Triple patterns derived from a BN's structure can also be used as input features to a classification tree. A graph of n nodes contains C_n^3 triples which can then be converted to a categorical vector of C_n^3 elements. In this study, n equals 7, resulting in 42-dimensional classification features for an EG and 35-dimensional features for exploring triple patterns.

The performance of BN-based classifiers were evaluated with both cross-subject validation and within-subject validation on a real sEMG data set containing repeated trials of arm reaching movements. In cross-subject validation (Fig. 4), all the trials of one arm side of one subject were kept aside as the testing data, and all the other data were used to train a classifier which was then used to predict the stroke state and hand dominance of the testing arm side. The stroke state and hand dominance of each testing trial was predicted individually, and then all the predictions voted on the state of the testing arm side. In this way, all the trials of the arm being tested were used to predict its group membership. This procedure is repeated for each subject in a leave-one-out, cross-validation manner. Cross-subject validation evaluates whether data of the same group share common features while data of different groups have distinguishing features. The strategy of within-subject validation is shown in Fig. 5. One trial of a subject was kept aside as the testing data, and all the other trials of the same subject were used to train a classifier which was then used to predict the group membership of the testing trial. This procedure was repeated and each time a different trial was selected as the testing trial. Within-subject validation was used to evaluate whether different trials of one arm from the same subject were consistent, yet trials of the other arm from the same subject were different.

The performance of the BN-based classification trees applied to sEMG carrier was compared with that of a PCA-based SVM applied to the sEMG amplitude [24]. The PCA-based SVM approach includes two steps: 1) reduce the dimension of sEMG amplitude with PCA. 2) input the dimension-reduced data to SVMs for classification. In our study, 6 principal components (PC) were needed to explain 80% of the total variance. The PC coefficients of the seven muscles in our study were then concatenated as input feature vectors to SVMs whose length was $6 \times 7 = 42$. We constrained both classification trees and SVMs from using no more than three elements of the input features to avoid over-fitting. We tried to input all the 35 or 42 features to the classifiers, but the corresponding classification error rates were as high as 30% to 50%, so we set the constraint above to improve the performance. All the combinations of no more than three features were exhaustively searched, and finally the best performance was selected. The results showed that this modified implementation provided much better performance than using all the features together. We think that the comparisons between the BN-based classification trees and the PCA-based SVMs are fair because: they were subject to the same constraint; the best combination of features were enumeratively searched; and the classification trees were not provided with more features than the SVMs were.

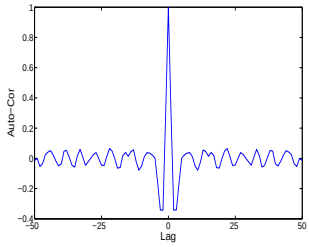


Fig. 6. A typical auto-correlation graph of the "carrier" signal $x(t)$ of a muscle. $x(t_1)$ and $x(t_2)$ ($t_1 \neq t_2$) are almost independent.

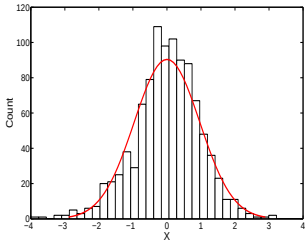


Fig. 7. A typical histogram of the time-distribution of the "carrier" signal $x(t)$ of a muscle. The distribution of $x(t)$ in the time domain is almost Gaussian.

E. Modeling sEMG signals

A sEMG signal $y(t)$ is usually considered as a zero-mean, Gaussian, band-limited and wide-sense stationary stochastic process $x(t)$ modulated by the EMG amplitude $a(t)$ [5], [25], expressed as

$$y(t) = x(t)a(t), \tag{11}$$

where t indicates time and $x(t)$ is named by us as "carrier", a term borrowed from the field of communication. Accepting these assumptions and being consistent with what we observed in this study, we assume

1) $x(t)$ is wide-sense stationary;
2) $x(t)$ follows a Gaussian process;
3) $x(t)$ is approximately white;
4) $x(t)$ is ergodic.

Fig. 6, a typical auto-correlation plot of $x(t)$, shows that $x(t)$ is approximately white, *i.e.* $x(t_1)$ and $x(t_2)$ are approximately independent if $t_1 \neq t_2$. Fig. 7, a typical histogram of the distribution of $x(t)$, shows that $x(t)$'s distribution is almost Gaussian in the time-domain. Since the distribution of $x(t)$ is suggested to be Gaussian both spatially and temporally, the ergodicity assumption at least is not severely violated yet not rigorously proved.

The four assumptions as a whole imply that a single-channel "carrier" signal $x(t)$ is independent identically-distributed (iid) at different time points and the distribution is well approximated as Gaussian. Therefore, we can model multi-channel "carrier" signals with static Gaussian BNs by adding one more assumption that the joint distribution of the multi-channel signals follows a multivariate Gaussian distribution. Since static models are used here, what we attempt to discover is not the dynamics of the muscles' activities, but the invariant interaction patterns among the muscles during the reaching movements.

Most of the existing literature on sEMG describes rectification and low-pass filtering of the data and hence emphasizes the amplitude $a(t)$ [5] while $x(t)$ the carrier is generally abandoned. However, in contrast here we focus on the "carrier" signals but not the amplitude, which will provide novel insights

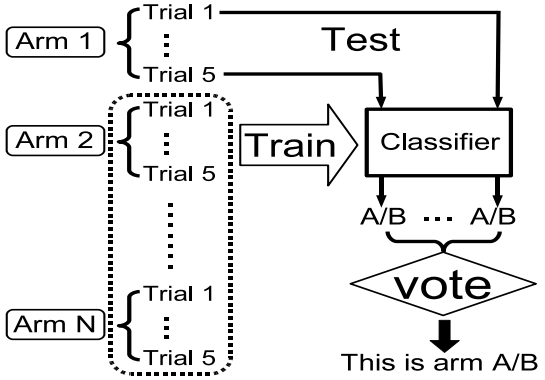


Fig. 4. Cross-subject validation

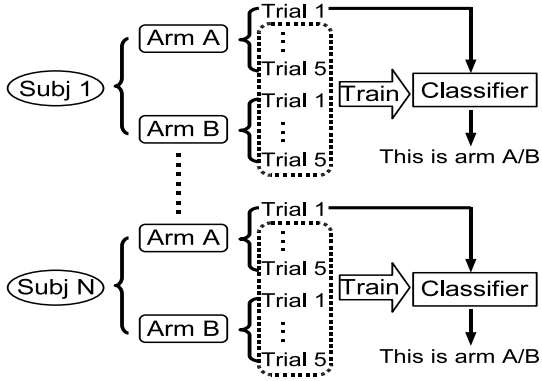


Fig. 5. Within-subject validation

into the underlying system. As supported by our analysis results reported in Sec. III, the "carrier" signal is also informative and provides a robust way to deal with the challenging issue of inter-subject variability in sEMG data.

III. Results

A. Real sEMG Datasets

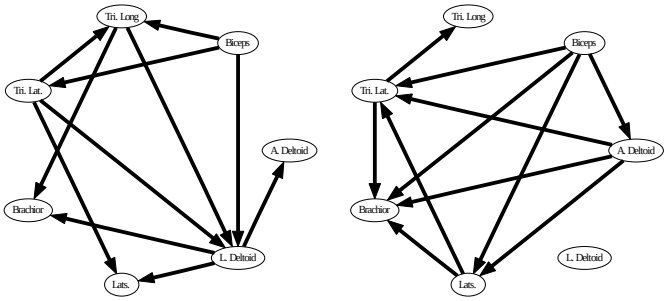
All research was approved by the University of British Columbia Ethics Board. Thirteen stroke subjects and 9 healthy subjects were recruited. In the experiment, subjects sat in a chair with their hands on the thigh, and then reached to a shoulder-height target as fast as they could for five to ten trials with each arm. The sEMG of the following seven muscles were collected: the deltoids (anterior and lateral), the triceps (long and lateral heads), the biceps brachium, the latissimus dorsi, and the brachioradialis. A bipolar montage was used to minimize the effect of crosstalk. The seven-channel sEMG signals were amplified, high-pass filtered at 20 Hz to reduce movement artifact, and then sampled at 600 Hz. (Please refer to [26] for further details on the sEMG experiment procedure). The amplitude of the sEMG was estimated with root-mean-square (RMS), with a moving window of 0.1 second. As EMG signal can be considered as a wide sense stationary stochastic process modulated by the EMG amplitude [5], the carrier stochastic process (see Sec. II-E) was also calculated by dividing the sEMG signal by the estimated amplitude. Finally, the sEMG signals (both the amplitude and carrier) of different trials were resampled with cubic spline interpolation so that the overall movement duration from reaching-start to target-touching was exactly 1000 time points. This prevents the sub-network analysis (see Sec. II-C) from being biased by the unequal data length since a DAG learned from more time points tend to include more connections than another learned from less time points. Preliminary studies included testing the effect of the moving window by stepwise increasing the width from 20 ms to 300 ms. It was determined visually that 100 ms gave the best estimation, as it yielded good amplitude estimate and produced an approximately wide-sense stationary carrier signal.

Since we were interested in the influence on sEMG patterns of two factors, stroke condition and hand dominance, sEMG recordings were grouped into four types of experimental groups: healthy dominant hand (HD), healthy non-dominant hand (HN), stroke more affected side involving the dominant hand (SD) and stroke more affected side involving the non-dominant hand (SN). To sharpen the contrast between the stroke and healthy states, the less affected side of stroke subjects was excluded because it may not be a valid comparison against the healthy state. Most individuals with stroke have subtle deficits on the non-paretic side due to a number of factors, including the contribution of the small portion of corticospinal tracts that do not decussate, and remain ipsilateral. To focus on the effect of one factor at a time, we fixed the state of one factor and compared the two states of the other factor in four group comparisons: HD *v.s.* HN, SD *v.s.* SN, HN *v.s.* SN and HD *v.s.* SD.

B. Learned Bayesian Networks

Examples of the learned DAGs of the BNs are given in Fig. 8, where the left side illustrates a typical DAG of the dominant hand side of a healthy subject (*i.e.* a HD case) and the right is for a typical DAG of the non-dominant hand side of a stroke subject (*i.e.* a SN case). The two DAGs showed different connection features. For example, the lateral deltoid is connected to all the other six muscles in the DAG of the HD case while it is completely isolated in the SN case. The long head of the triceps is more connected with others in the HD case than in the SN case. The biceps is connected to the lateral, long heads of the triceps and the lateral deltoid in the HD case while to the anterior deltoid, latissimus dorsi, the brachioradialis and the lateral triceps in the SN case. These differences between the two typical subjects suggest that both stroke and hand dominance can affect the sEMG muscle association patterns during reaching movements.

Comparisons between the DAGs of different hand dominance are shown in Fig. 9. The width of the displayed edge is proportional to the log odds ratio (LOR) of their appearance rates in the two groups. First the DAGs were converted to EGs, then the appearance rate of each edge in each group of the EGs was calculated, and finally the rate was converted to LOR. Only the edges whose LOR's absolute value exceeded ln(2) were shown in the figure. A solid edge means that it appears more frequently in the first group than in the second, and a dashed edge means a higher appearing frequency in the second



(a) Healthy, Dominant (b) Stroke, Non-dominant

Fig. 8. Examples of typical DAGs of subjects with different hand dominance and different stroke state.

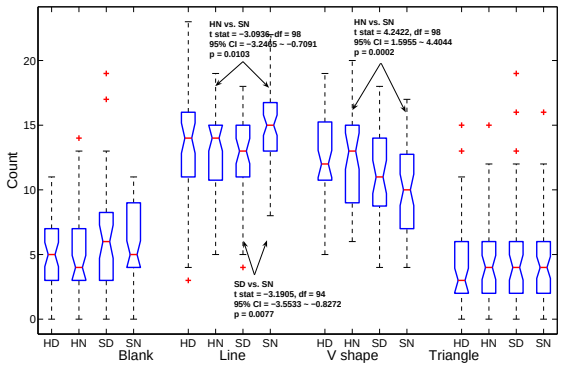


Fig. 10. Comparison of the count of appearances of connection patterns in different types of experimental groups. The distribution of the number of appearances is shown with box plot. The boxes have lines at the lower quantile, median and the higher quantile. The whiskers are the extent of the rest of the data and the plus symbols are the outliers. If the notches of two boxes overlap, their medians differ significantly with type I error rate less than 5%. The means of the distributions are also compared with t-test and significant results are labeled with arrows. p-values are adjusted for multiple comparisons with Sidak correction. Since the notches compare the medians and the t-test compares the means, their results may differ from each other when the distributions are skewed, for example in the comparison of pattern "Line" between HN and SN. There are no statistical differences in the number of edges between the above groups, and thus the statistically different results between groups are not based on the number of edges.

group. Sub-figures (a) and (b) compared the BNs learned from dominant hand and non-dominant hand groups. We note that networks from dominant hand groups have more connections between the muscle pairs (biceps, triceps long head), (lateral triceps, lateral deltoid) and (anterior deltoid, lateral deltoid). Sub-figures (c) and (d) compared the BNs between healthy subjects and stroke subjects. We note that healthy subject group has more connections between the muscle pairs (triceps long head, biceps) and (brachioradialis, lateral deltoid).

C. Triple Patterns

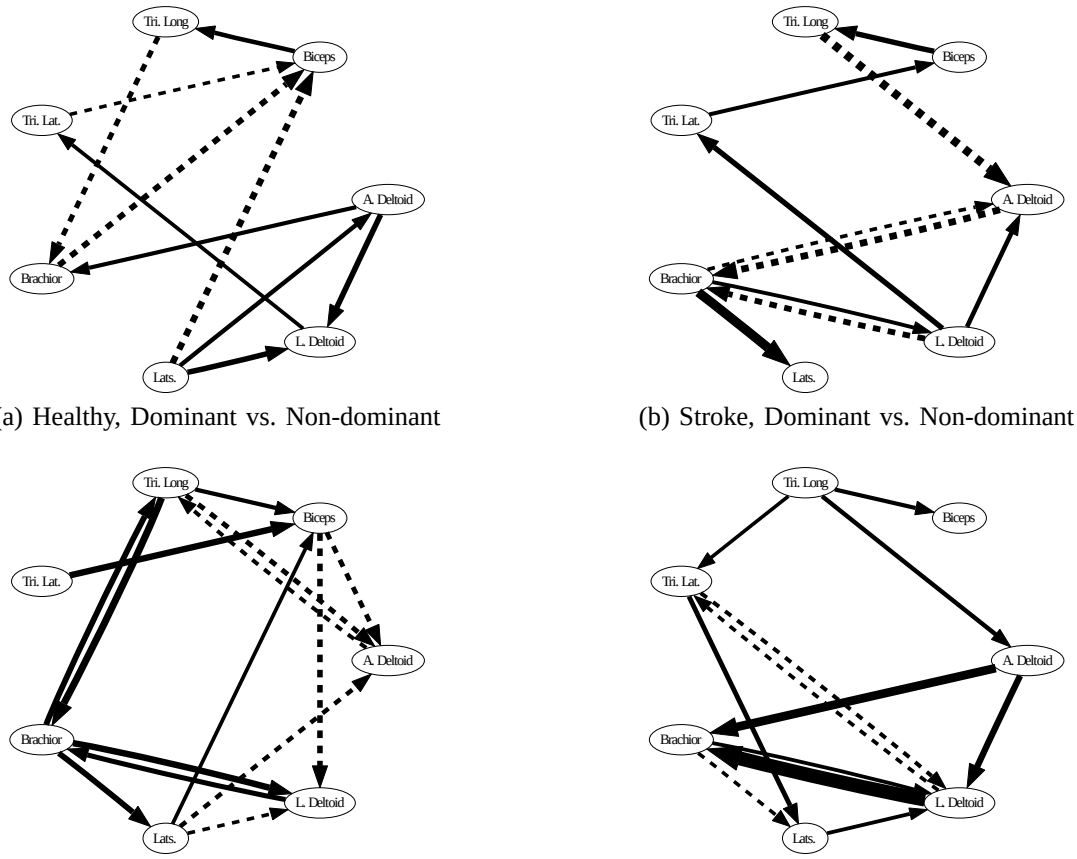
The learned BNs from different experimental groups demonstrated different triple connection patterns, as shown in Fig. 10. The V pattern (see Fig. 3) appears significantly more frequently in the HN group than in the SN group ($p = 0.0002$), and more often in the SD group than in the SN group. The Line pattern (see Fig. 3) appears significantly more frequently in the SN group than in the HN group ($p = 0.0103$), and also in the SN group than in the SD group ($p = 0.0077$). In addition to the general triple patterns, specific muscles triples also showed significantly different connection patterns across experiment groups. Muscles involved in these triplets are deltoid (both anterior and lateral), triceps and brachioradilis, as shown in Table I. However, no significant results were discovered about the Blank pattern and the Triangle pattern (see Fig. 3).

Since the V pattern is the most efficient pattern to connect three muscles and a Line pattern only connects two muscles, muscles of the HN and SD groups seem cooperate more closely than those of the SN group. This observation coincides with clinic experience, where the SN group typically have the most difficulty in performing reaching movements. Lack of normal cooperation between the muscles may explain this empirical observation of the SN group's demonstrating inferior performance [27]. The importance of the deltoid (both anterior and lateral) in these results is consistent with a previous traditional analysis of these data [26] where the deltoids' activation was found significantly altered after stroke. The functional connectivity between the brachioradilis and the deltoid that we detected (Fig. 9 c-d) during reaching movements has been suggested by previous studies. Lemon et al. [28] used transcranial magnetic stimulation during reaching movements in human subjects and found evidence of a strong cortical drive to both the deltoid and brachioradialis throughout a reaching movement. The connectivity between the deltoid and triceps found more prominently in the stroke subjects (Fig. 9 c-d) may suggest a more traditional stroke synergy, where there is breakdown in the normal independent activation of muscles involving the shoulder girdle and those involved in movement of the elbow [29]. Probably because the reaching task is sufficiently simple that healthy subjects mastered it easily even with their non-dominant hands, no significant difference between HD and HN groups was found. An alternative explanation is that handedness may not be strongly contrasted in individuals that exhibit forms of ambidexterity.

D. Classification Performance

The across-subject classification performances were reported in Table II. The proposed methods, which use the structure features of the BNs learned from the carrier signals, provide very high classification accuracy in the four classification tasks, and outperform the PCA-based SVM approach applied to the amplitude signals. This excellent classification performance is unlikely related to over-fitting because the error rate was estimated with cross-subject validation and the classifiers were provided with almost equal chances to achieve good performance. Although most of current studies on sEMG focus on the amplitude signals for classification, our classification trees are based on the "carrier" signal (see Sec. II-E) which is usually lost in the traditional process of rectification, smoothing and other preprocessing steps. In addition to almost perfect classifications involving stroke, the proposed methods also offer visible interpretations via way of the classification tree.

The BN-based classification trees generally provided better performance when the triple patterns were used as the input



(a) Healthy, Dominant vs. Non-dominant

(b) Stroke, Dominant vs. Non-dominant

(c) Non-dominant, Healthy vs. Stroke

(d) Dominant, Healthy vs. Stroke

Fig. 9. Mean differences of overall network structures as a function of condition. The labels under the sub-figures are the experimental conditions in comparison. A solid edge means it appears more frequently in the first type than in the second, and a dashed edge does vice versa. The width of an edge is proportional to the contrast of the frequencies of its appearances. The contrast are measured with log odds ratio (LOR) which is defined as LOR = $\ln([f_1/(1-f_1)]/[f_2/(1-f_2)])$ where f_1 and f_2 are the appearance frequencies of the edge in two types of experimental situations under contrast. If the edge appears k times in n trials, the appearance frequency f is estimated as $k+1/(n+2)$ with the Bayes estimator. The Bayes estimator is more robust than MLE when k and n are small, and converges to MLE when k and n are large. Only edges whose absolute value of the LOR is greater than ln(2) are shown in the figure.

TABLE I
SIGNIFICANT CONNECTION PATTERNS OF SPECIFIC TRIPLES.

Comparison	Triple	Pattern	Count	OR and 95% CI	p-value	q-value
HD vs. SD	A. Deltoid, L. Deltoid and Tri. Lat.	L	13/45 vs. 0/41	inf, (3.5052, inf)	0.0151	0.0076
	L. Deltoid, Tri. Lat. and Tri. Long	L	13/45 vs. 0/41	inf, (3.5052, inf)	0.0151	0.0076
HN vs. SN	Brachior, L. Deltoid and Tri. Lat.	L	10/45 vs. 34/55	0.1765, (0.0650, 0.4638)	0.0140	0.0141
SD vs. SN	A. Deltoid, Brachior and Tri. Long	B	24/41 vs. 9/55	7.2157, (2.5614, 21.0100)	0.0036	0.0036
	A. Deltoid, L. Deltoid and Tri. Lat.	L	0/41 vs. 14/55	0.0000, (0.0000, 0.3312)	0.0390	0.0157

The connection patterns of the specific muscle triples appears significantly more/less frequently in a type of experimental groups than in another type. Counts are in the form of (No. of appearance / No. of trials). OR and CI are short for "odds ratio" and "confidence interval". p-values are originally calculated with Fisher's exact test and are then adjusted for multiple comparisons with the Sidak correction or converted to q-values with the FDR [22].

classification features than when the adjacency matrices of EGs were used. We think this is because an EG just encodes pair-wise interactions but a triple involves three muscles. Figs. 11 and 12 showed the best classification trees for the four classification tasks. These trees include interactions between agonist-antagonist pairs (*e.g.* long head of triceps, biceps), muscles with similar actions (*e.g.* biceps, brachioradialis), and muscles with no obvious similarity of function but might be part of larger synergies (*e.g.* brachioradialis and latissimus dorsi).

As previously mentioned in Sec. II-D, in addition to using their structure features, BNs by themselves can be extended straightforwardly to a statistical classifier. The BN statistical classifier separated the trials of the same subject perfectly (*i.e.* the within-subject cross-validation error rate = 0%) but performed poorly in cross-subject validation (*i.e.* the error rate $\approx$ 50%). Fig. 13 demonstrated the usage of a BN classifier for classifying healthy subjects' dominant and non-dominant hands. Although the BN classifier showed high trial-to-trial reliability, its across-subject classification performance was poor. We believe that the poor across-subject performance of the BN statistical classifiers is not due to deficiencies in the method, but rather reflective of the underlying biology. We note that the likelihood function is very consistent across

TABLE II
THE ERROR RATES OF CROSS-SUBJECT CLASSIFICATIONS

Signal Feature Classifier	Car. EG CT	Car. Triple CT	Amp. PCs SVM
HD vs. HN	**0.0556**	0.1111	0.1667
SD vs. SN	**0.0000**	**0.0000**	0.0769
HN vs. SN	**0.0000**	**0.0000**	0.1250
HD vs. SD	0.1333	**0.0000**	0.0667

"Car." and "Amp." are short for the carrier and the amplitude respectively. "EG" and "Triple" are the essential graphs and the triple patterns of the carrier's Bayesian networks. "PC" is short for principal component. "CT" and "SVM" are short for classification trees and support vector machines respectively. All the error rates were estimated with the cross-subject validation procedure in Fig. 4. Bold numbers are the best performance of the four classifiers.

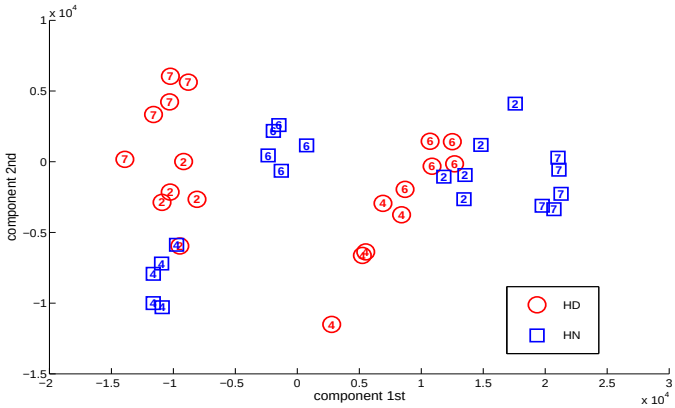


Fig. 13. Principal components of log likelihood of BNs trained from sEMG data of the dominant (HD) and non-dominant (HN) arms in healthy subjects. As demonstrated in Eqs. (9) and (10), a statistical BN classifier is a function of $p(x|M)$ where x is the data of a trial and M is the BN model of another trial. Thus, a trial is represented as a log likelihood vector composed of $p(x|M_i)$ where M_is are the models of many trials. To visualize the high dimension vectors, we plot their first two principal components. Note that trials of the same arm tend to cluster together, which suggests that the BN represents the reaching movement reliably and consistently. Trials of HD and HN are not separated, which suggests that arms of the same type do not share a common distinguishing pattern in their log likelihood, but each has a different pattern. Since a statistical BN classifier is based on log-likelihood, its cross-subject performance is poor while its within-subject performance is excellent.

trials of the same subject, which suggests that it is robust to various artifacts that may corrupt sEMG signals. Nevertheless, there may be considerable variations between individuals due to factors such as variance in genetics, developmental environment, compensatory strategies, and ongoing plasticity in response to environmental stressors. Thus a key result of the present work is implication that most within-group, inter-subject variability is not in the network structure (which itself is sensitive to handedness and effects of stroke), but rather in the parameters that specify the interactions within the network structure.

IV. Conclusions and Discussions

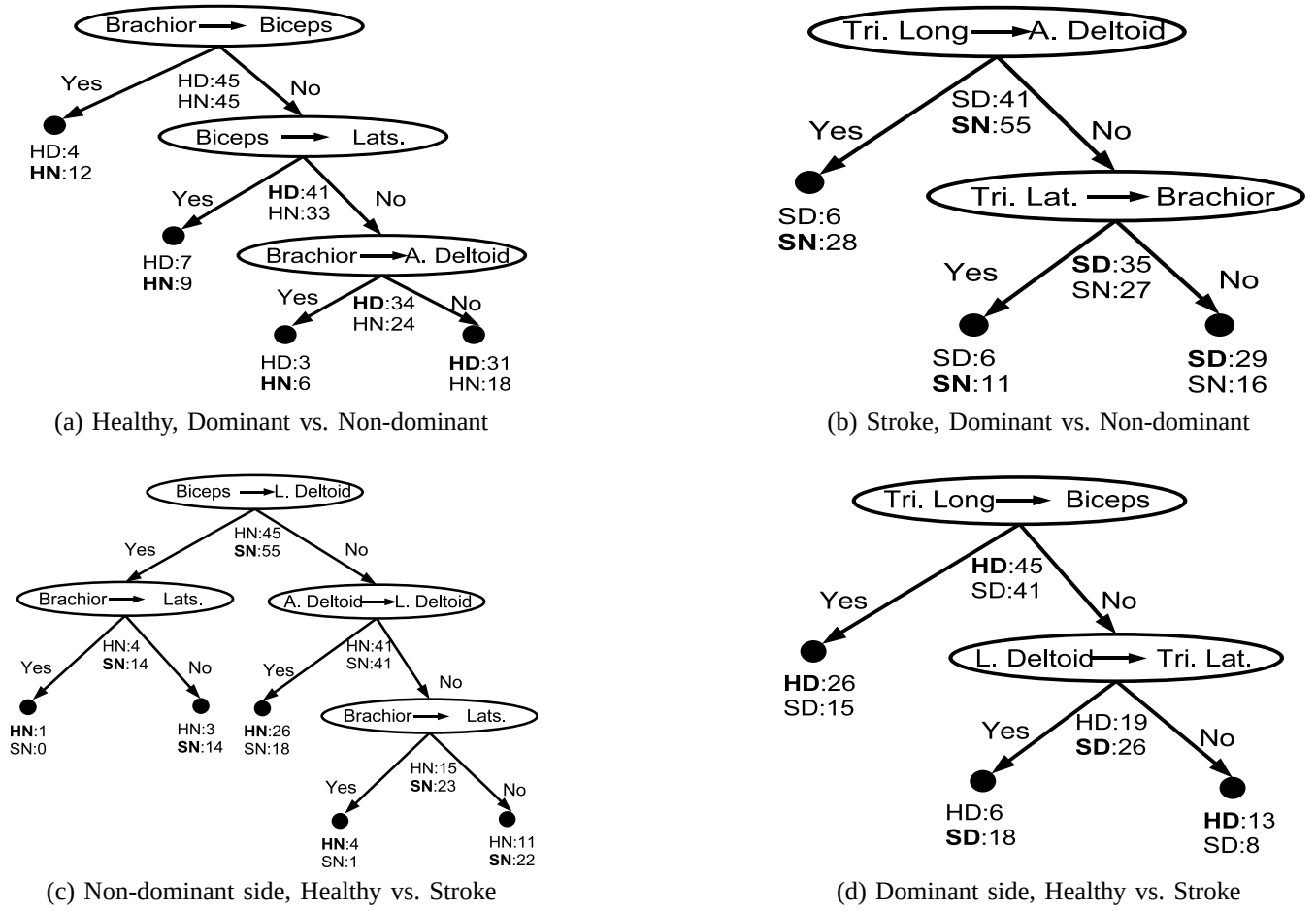
In this paper, we have developed a Bayesian network (BN) framework based for modeling muscle networks to represent muscle co-ordinations in motor control. To demonstrate the benefits of the proposed approach, we applied this method to multi-channel sEMG data simultaneously recorded during reaching movements in healthy and stroke subjects. We noted that dependent muscle synergies can be revealed by first using a BN to model the muscle interaction network and then analyzing subnets of the derived BN. Further examination of the muscle synergies (muscle association patterns) suggested that stroke may particularly affect the interaction between a few specific muscles, especially the deltoid, during reaching movements. Classification trees based on the BNs' structure features can effectively classify the reaching performed by the healthy dominant, healthy non-dominant, stroke dominant and stroke non-dominant arms across subjects. Classification trees provide the additional benefit of providing a visible means to view the classification.

A key result of this study was that statistically significant differences between the sEMG recorded under different conditions were noticed from analyzing the "carrier" signal, which is obtained by estimating and removing the amplitude information from the raw sEMG data. This is noteworthy observation, since in many sEMG studies, only the amplitude is investigated with the carrier signal discarded. Although not specifically explored in this study, we believe that the widespread statistical dependencies between the "carrier" of the sEMG signals from different muscles may reflect widespread synchronization between different cortical areas and muscles known to exist during dynamic movements [30] [7]. We are intrigued that the significant features of the networks (Fig. 9 c-d) suggest a significant statistical interaction between the deltoid and brachioradialis, consistent with previously-described *cortical-muscle* interactions during reaching movements in normal subjects [28].

Another promising result from this study is the finding that the *structure* of BNs and their subsets are quite robust across individuals within the same group, yet demonstrate enough sensitivity to detect handedness and the effects of stroke. One of the fundamental challenges in classification of movements after stroke is how to deal with the inherent subject-to-subject variability in sEMG recordings, yet still be sensitive enough to detect impairment. The overall BN models demonstrated robustness to trial-to-trial variability within subjects (Fig. 13), but were quite different across subjects within the same group. However, the success of the classification trees based on the *structure features* of the subject-specific BNs suggests that the BN structure is sensitive to the effects due to stroke or handedness factor, but robust to inter-subject variability.

Although we suggest that our results demonstrate strong evidence to support the use of BNs as a tool to study sEMG signals, there are nevertheless shortcomings of the proposed method. The BN model, as proposed here, assumes stationarity of the muscle interactions. Yet there is evidence that muscle interactions may be dynamically affected by a number of factors. For example, different heads of the gastrocnemius muscle may be activated during walking as a function of activity in sensory afferents [31], an observation with neuroanatomical basis [32]. Presumably integrating the time-varying amplitude information with a dynamic Bayesian network (DBN) in addition to the carrier information for classification of sEMG data may be a fruitful avenue to explore in the future.

Each sEMG channel is a measure of the depolarization of muscle fibers, which is the end result of motor cortex activity,



(a) Healthy, Dominant vs. Non-dominant

(b) Stroke, Dominant vs. Non-dominant

(c) Non-dominant side, Healthy vs. Stroke

(d) Dominant side, Healthy vs. Stroke

Fig. 11. The best classification trees by using EG structures as the classification features. If two trees have equal cross-subject classification error rates, the one using fewer input edges is chosen. For each node, a label specifies its group membership and its predicted value (in bold). For a branch, its label is the decision rule. For the performances of these trees, please refer to Table II.

conduction along peripheral nerves, propagation across the neuromuscular junction, and propagation within the muscles. At each stage of motor propagation, there is the possibility for temporal variability, for example, "neuromotor noise" in the central nervous system which may be particularly important in disease states, conduction velocity along the peripheral nerves, which is a strong function of temperature, jitter at the neuromuscular junction and possible disease states in the muscles themselves [33], [34]. However, successfully incorporating these features in the future will additionally involve explicitly modeling the dynamic temporal relationships between sEMG recordings.

We do not suggest that the model is a biologically-accurate description of the generative process of EMG activity. Nevertheless, we note that both Renshaw cells and Ia inhibitory interneurons have complex effects on the firing of α motor neurons in the spinal cord. Since these modulatory cells are themselves activated by muscle activities [9], the "dependent synergies" proposed – where interactions between muscles are influenced by the activity of other muscle(s) – are physiological plausible. However, explicitly incorporating inhibitory interneurons would require the addition of hidden nodes, beyond the scope of the current proposed method.

ACKNOWLEDGMENT

This research was partially funded by a National Parkinsons Foundation (NPF) Center of Excellence Grant and a Natural Sciences and Engineering Research Council of Canada (NSERC) Grant.

REFERENCES

[1] A. d'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature neuroscience*, vol. 6, pp. 300–308, 2003.

[2] M. C. Tresch1, P. Saltiel1, and E. Bizzi1, "The construction of movement by the spinal cord," *Nature Neuroscience*, vol. 2, no. 2, pp. 162–7, 1999.

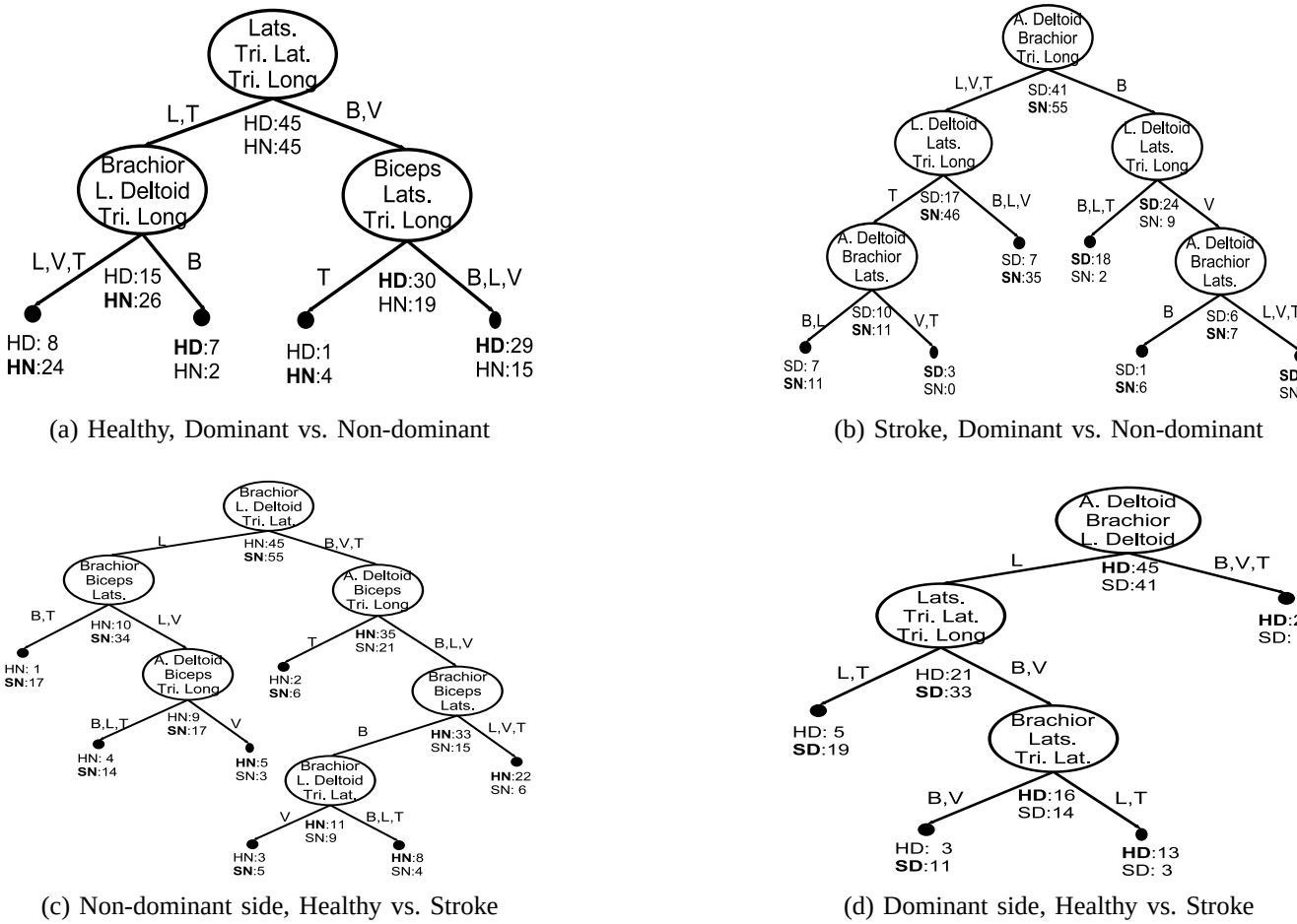
[3] R. Merletti, D. Farina, and A. Granata, "Non-invasive assessment of motor unit properties with linear electrode arrays," *Electroencephalography and clinical neurophysiology. Supplement*, vol. 50, pp. 293–300, 1999.

[4] N. Stoykov, M. Lowery, A. Taflove, and T. Kuiken, "Frequency- and time-domain FEM models of EMG: capacitive effects and aspects of dispersion," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 8, pp. 763–72, Aug. 2002.

[5] E. A. Clancy, E. L. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *Journal of Electromyography and Kinesiology*, vol. 12, no. 1, pp. 1–16, Feb. 2002.

[6] M. J. McKeown, "Cortical activation related to arm-movement combinations," *Muscle & Nerve*, vol. 23, no. S9, pp. S19–S25, 2000.

[7] M. McKeown and R. Radtke, "Phasic and tonic coupling between EEG & EMG revealed with independent component analysis (ICA)," *J Clin Neurophysiology*, vol. 18, pp. 45–57, 2001.



(a) Healthy, Dominant vs. Non-dominant

(b) Stroke, Dominant vs. Non-dominant

(c) Non-dominant side, Healthy vs. Stroke

(d) Dominant side, Healthy vs. Stroke

Fig. 12. The best classification trees by using triple patterns as the classification features. If two trees have equal cross-subject classification error rates, the one using fewer input edges is chosen. For each node, a label specifies its group membership and its predicted value (in bold). For a branch, its label is the decision rule. For the performances of these trees, please refer to Table II.

[8] M. C. Tresch, V. C. K. Cheung, and A. d'Avella, "Matrix factorization algorithms for the identification of muscle synergies: evaluation on simulated and experimental data sets," *J Neurophysiol*, vol. 95, pp. 2199–2212, 2006.

[9] S. Rossignol, R. Dubuc, and J.-P. Gossard, "Dynamic sensorimotor interactions in locomotion," *Physiol. Rev.*, vol. 86, no. 1, pp. 89–154, 2006.

[10] T. J. Carroll, E. R. L. Baldwin, and D. F. Collins, "Task dependent gain regulation of spinal circuits projecting to the human flexor carpi radialis," *Experimental Brain Research*, vol. V161, no. 3, pp. 299–306, Mar. 2005.

[11] M. Kaminski, "Determination of transmission patterns in multichannel data," *Phil. Trans. R. Soc. B*, vol. 360, no. 1457, pp. 947–952, May 2005.

[12] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: Simple building blocks of complex networks," *Science*, vol. 298, pp. 824–827, 2002.

[13] K. A. Provins, "Handedness and speech: A critical reappraisal of the role of genetic and environmental factors in the cerebral lateralization of function," *Psychological Review*, vol. 104, no. 3, pp. 554–571, Jul. 1997.

[14] R. L. Sainburg, "Evidence for a dynamic-dominance hypothesis of handedness," *Experimental Brain Research*, vol. 142, no. 2, pp. 241–258, Jan. 2002.

[15] K. P. Murphy, "Dynamic Bayesian networks: representation, inference and learning," Ph.D. dissertation, University of California, Berkeley, 2002, chair-Stuart Russell.

[16] S. L. Lauritzen, *Graphical models*. Oxford, New York: Clarendon Press, Oxford University Press, 1996, vol. 17.

[17] S. A. Andersson, D. Madigan, and M. D. Perlman, "A characterization of Markov equivalence classes for acyclic digraphs," *The Annals of Statistics*, vol. 25, pp. 505–541, 1997.

[18] N. Friedman and D. Koller, "Being Bayesian about network structure. a Bayesian approach to structure discovery in Bayesian networks," *Machine Learning*, vol. V50, no. 1, pp. 95–125, Jan. 2003.

[19] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, Mar. 1978, gR: Short Communications.

[20] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[21] K. Murphy, "Bayes net toolbox for Matlab (BNT)." [Online]. Available: http://bnt.sourceforge.net/

[22] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society.Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[23] V. N. Vapnik, *The Nature of Statitical Learning Theory*. Springer-Verlag, 1995.

[24] N. F. Guler and S. Kocer, "Classification of EMG signals using PCA and FFT," *Journal of Medical Systems*, vol. V29, no. 3, pp. 241–250, Jun. 2005.

[25] E. A. Clancy, "Electromyogram amplitude estimation with adaptive smoothing window length," *Biomedical Engineering, IEEE Transactions on*, vol. 46, no. 6, pp. 717–729, 1999.

[26] P. H. McCrea, J. J. Eng, and A. J. Hodgson, "Saturated Muscle Activation Contributes to Compensatory Reaching Strategies After Stroke," *J Neurophysiol*, vol. 94, no. 5, pp. 2999–3008, 2005.

[27] J. Harris and J. Eng, "Individuals with the dominant hand affected following stroke demonstrate less impairment than those with the non-dominant hand affected," *Neurorehabilitation and Neural Repair*, in press.

[28] R. Lemon, R. Johansson, and G. Westling, "Modulation of corticospinal influence over hand muscles during gripping tasks in man and monkey,"

Canadian Journal of Physiology and Pharmacology, vol. 74, pp. 547–558, 1996.

[29] P. L. Gribble and D. J. Ostry, "Independent coactivation of shoulder and elbow muscles," *Experimental Brain Research*, vol. V123, no. 3, pp. 355–360, Nov. 1998.

[30] B. Feige, A. Aertsen, and R. Kristeva-Feige, "Dynamic Synchronization Between Multiple Cortical Motor Areas and Muscle Activity in Phasic Voluntary Movements," *J Neurophysiol*, vol. 84, no. 5, pp. 2622–2629, 2000.

[31] J. Duysens, B. M. H. van Wezel, T. Prokop, and W. Berger, "Medial gastrocnemius is more activated than lateral gastrocnemius in sural nerve induced reflexes during human gait," *Brain Research*, vol. 727, no. 1-2, pp. 230–232, Jul. 1996.

[32] L. A. LaBella, J. P. Kehler, and D. A. McCrea, "A differential synaptic input to the motor nuclei of triceps surae from the caudal and lateral cutaneous sural nerves," *J Neurophysiol*, vol. 61, no. 2, pp. 291–301, 1989.

[33] D. J. Reinkensmeyer, M. G. Iobbi, L. E. Kahn, D. G. Kamper, and C. D. Takahashi, "Modeling reaching impairment after stroke using a population vector model of movement control that incorporates neural firing-rate variability," *Neural Comp.*, vol. 15, no. 11, pp. 2619–2642, 2003.

[34] K. R. Mills, "Specialised electromyography and nerve conduction studies," *J Neurol Neurosurg Psychiatry*, vol. 76, no. suppl 2, pp. ii36–40, 2005.

Junning Li received his BSc in Automation from Tsinghua University, China, in 2002, and subsequently MSc in Control Science and Engineering in 2005 also from Tsinghua University. He currently is a Ph.D candidate at the department of Electrical and Computer Engineering of the University of British Columbia, Canada, working on his thesis "Dynamic Bayesian Networks: Modelling and Analysis of Neural Signals".

Z. Jane Wang received the B.Sc. degree from Tsinghua University, China, in 1996, with the highest honor, and the M.Sc. and Ph.D. degrees from the University of Connecticut in 2000 and 2002, respectively, all in electrical engineering. While at the University of Connecticut, Dr. Wang received the Outstanding Engineering Doctoral Student Award. She has been Research Associate of the Institute for Systems Research at the University of Maryland, College Park. Since Aug. 1, 2004, she is with the Department Electrical and Computer Engineering at the University of British Columbia (UBC), Canada, as an Assistant Professor. Her research interests are in the broad areas of statistical signal processing, with applications to information security, biomedical imaging, genomic and bioinformatics, and wireless communications. She co-received the EURASIP Best Paper Award 2004 and IEEE Signal Processing Society Best Paper Award 2005, and a Junior Early Career Scholar Award from Peter Wall Institute at UBC in 2005. She co-edited a book Genomic Signal Processing and Statistics and co-authored a book Multimedia Fingerprinting Forensics for Traitor. She is an Associate Editor for the EURASIP Journal on Bioinformatics and Systems Biology.

Janice J. Eng received her BSc in Physical Therapy and Occupational Therapy in 1985 from the University of British Columbia, Canada. She received a MSc in Biomedical Engineering in 1987 from the University of Toronto and PhD in Kinesiology in 1994 from the University of Waterloo, Canada. She completed post-doctoral work in neurophysiology at the Simon Fraser University, Canada. She is Professor, School of Rehabilitation Sciences, at the University of British Columbia and the GF Strong Rehab Centre. She currently holds career scientist awards from the Michael Smith Foundation of Health Research and the Canadian Institutes of Health Research. She is on the Editorial Board for the journal, Physical Therapy. Her research applies principles of rehabilitation, neurophysiology, exercise physiology, and mechanics to investigate the effects of neurological impairments and their rehabilitation treatments on postural control and mobility.

Martin J. McKeown graduated *summa cum laude* in Engineering Physics from McMaster University in 1986, and subsequently earned his MD degree from the University of Toronto in 1990. From 1990-1994 he specialized in neurology at the University of Western Ontario, and later did a fellowship in clinical electrophysiology. He is Board Certified in neurology in Canada and the US, and has been an examiner for the US Neurology Board certification. From 1995-98, he was a research associate at the Computational Neurobiology Lab, Salk Institute for Biology Studies, San Diego. From 1998-2003 he was an assistant professor of Medicine at Duke University, core faculty at Dukes Brain Research and Analysis Center, as well as an adjunct professor in Biomedical Engineering and Dukes Center for Neural Analysis. He is currently an Associate Professor of Medicine (Neurology) and faculty member of the Pacific Parkinsons Research Center, and the Brain Research Centre at the University of British Columbia. His current work includes electrophysiology and fMRI as it pertains to patients with movement disorders. In addition to research, he trains physicians in neurology and treats patients with Movement Disorders, such as Parkinsons disease.