Biomedical Signal Processing and Control 13 (2014) 327–336

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Surface myoelectric signal classification using the AR-GARCH model

Ghulam Rasool [b,*], Nidhal Bouaynaya [a], Kamran Iqbal [b], Gannon White [c]

[a] Department of Electrical and Computer Engineering, Rowan University, Glassboro, NJ 08028, USA
[b] Systems Engineering Department, University of Arkansas at Little Rock, Little Rock, AR 72204, USA
[c] Department of Health, Human Performance, and Sport Management, University of Arkansas at Little Rock, Little Rock, AR 72204, USA

ARTICLE INFO

Article history:
Received 4 April 2014
Received in revised form 23 May 2014
Accepted 2 June 2014
Available online 5 July 2014

Keywords:
Myoelectric signal
Autoregressive (AR) model
Heteroscedasticity
Autoregressive-autoregressive generalized conditional heteroscedastic (AR-GARCH) model
Myoelectric control

ABSTRACT

In myoelectric prostheses design, it is normally assumed that the necessary control information can be extracted from the surface myoelectric signals. In the pattern classification paradigm for controlling myoelectric prosthesis, the autoregressive (AR) model coefficients are generally considered an efficient and robust feature set. However, no formal statistical methodologies or tests are reported in the literature to analyze and model the myoelectric signal as an AR process. We analyzed the myoelectric signal as a stochastic time-series and found that the signal is *heteroscedastic*, i.e., the AR modeling residuals exhibit a time-varying variance. Heteroscedasticity is a major concern in statistical modeling because it can invalidate statistical tests of significance which may assume that the modeling errors are uncorrelated and that the error variances do not vary with the effects being modeled. We subsequently proposed to model the myoelectric signal as an autoregressive-generalized autoregressive conditional heteroscedastic (AR-GARCH) process and used the model parameters as a feature set for signal classification. Multiple statistical tests including the Ljung–Box Q-test, Engle's test for heteroscedasticity, the Kolmogorov–Smirnov test and the *goodness of fit* test were performed to show the validity of the proposed model. Our experimental results show that the proposed AR-GARCH model coefficients, when used as a feature set in two different classification schemes, significantly outperformed ($p < .01$) the conventional AR model coefficients.

1. Introduction

The pattern classification for myoelectric control is based on the assumption that the muscle activations are distinguishable and repeatable. The direct classification of the myoelectric signal is not preferred due to its large dimensionality and stochastic nature [1]. Therefore, the myoelectric signal is first segmented to form *analysis windows*. A set of representative features is extracted from each analysis window and is subsequently fed to a classifier [2,3]. Depending on the type of the classification algorithm, the classifier may require supervised or semi-unsupervised training [4,5]. Once trained, the output of the classification algorithm is a single identified movement class based on the input myoelectric signal. Various classification algorithms are employed for the purpose of myoelectric signal classification, e.g., the linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), different flavors of artificial neural networks (ANN), k-nearest neighbor (k-NN), neuro (fuzzy) schemes, Gaussian mixture models (GMM), and hidden Markov models (HMM) [2,3,5] (and references therein). Advanced classifiers including the support vector machines (SVM) [6], twin SVM [7] and self-enhancing LDA and QDA [8] have also been recently proposed. Furthermore, blind source separation (BSS) techniques including the independent component analysis (ICA) are also used to classify myoelectric signals [9,10]. Generally the pattern classification systems have produced promising results with classification accuracies in the range of 90% or more [2,3].

In the myoelectric control, the choice of an appropriate feature set is crucial for the pattern classification. Parker et al. have shown that the classification accuracy in a pattern classification scheme is more affected by the choice of the feature set than by the classification algorithm [11]. A common and promising feature set considered in the myoelectric control literature consists of autoregressive (AR) coefficients [2,5,8]. Tkach et al. investigated different features under conditions of change in muscle effort level, muscle fatigue and electrode shift, and found the AR coefficients to be one of the most robust features [12]. Similarly, Young et al. found that, under conditions of socket misalignment, the AR features performed better than the time-domain features [13]. However, in the myoelectric control literature, there has not been an exclusive effort to model the myoelectric signal as an AR

* Corresponding author. Tel.: +1 501 319 2167.
E-mail address: gxrasool@ualr.edu (G. Rasool).

http://dx.doi.org/10.1016/j.bspc.2014.06.001
1746-8094/

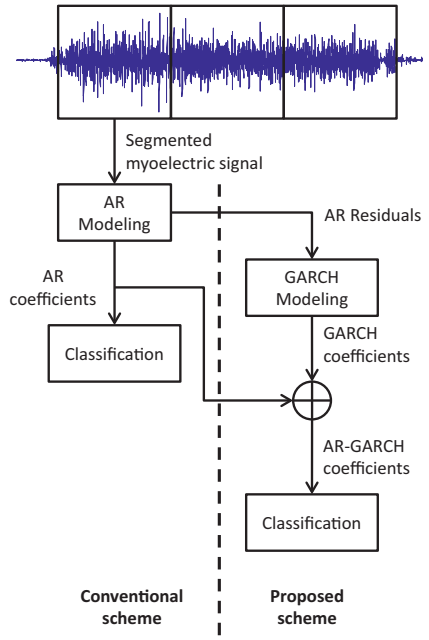

Fig. 1. Proposed and conventional approaches to AR modeling of the myoelectric signal. The proposed approach captures more information from the myoelectric signals as the AR residuals are further modeled as the GARCH process.

process using formal statistical tools. A generally adopted procedure in the literature is to adjust (increase or decrease) the AR model order based on the required classification accuracy and available computational resources. Such practice usually results in an AR model order in the range of 4–6 [2,5]. We believe that a detailed statistical study is essential to, (1) identify the parsimonious AR model order for the myoelectric signal, (2) measure the *goodness of fit* of the identified model, and (3) evaluate various characteristics of the modeling residuals, such as the sample autocorrelation and/or the heteroscedasticity. In the present work, we show that modeling the myoelectric signal as an AR process results into the residual errors that exhibit *heteroscedasticity*, i.e., "variability" in the variance. [14]. Concept of heteroscedasticity, autoregressive conditional heteroscedasticity (ARCH) and generalized ARCH (GARCH) models are introduced in Appendix A.

In biomedical engineering applications, the GARCH models have been employed to extract features from the electroencephalogram (EEG) signals within the framework of wavelet transforms [15]. The GARCH models have also been used within a state-space framework and the Kalman filter for modeling non-stationary variance in the EEG signal [16], or modeling covariance for generation of the EEG signal [17]. In the case of myoelectric signals, the GARCH process has been used for noise suppression with wavelet coefficients [18]. However, to the best of our knowledge, no exclusive effort is reported in the literature to model the myoelectric signal as an AR-GARCH process and use its parameters as a basis for statistical inference and biophysical interpretation. In Fig 1, we contrast the proposed approach with the conventional AR feature approach. The proposed AR-GARCH model extracts more information from the myoelectric signal as compared to the conventional scheme as the AR residuals further modeled as a GARCH process.

2. Material and methods

The study received approval from the Institutional Review Board at the University of Arkansas at Little Rock. All participants provided written informed consent before start of the data collection. Experimental protocol for data collection was designed as per the recommendations of surface electromyography for the non-invasive assessment of muscles (SENIAM) project [19] and standard physiology and anatomy text [20].

2.1. Participants

It has been shown that there is a significant positive correlation between myoelectric data collected from able-bodied participants and amputees [21]. Therefore a set of healthy participants can form a reasonable basis for testing newly proposed feature set. A total of six able-bodied participants, male and female volunteered for the study. All selected participants were right hand dominant with no neuromuscular disorder history.

2.2. Movements

We identified eleven movements for our data collection, i.e., five hand movements and six wrist movements. Hand movements included lateral grasp, cylindrical grasp, tripod grasp, hand open and index finger point [5]. The wrist movements included pronation and supination, ulnar and radial deviation, and flexion and extension. For the classification purpose a "no movement" class was also added, thus making 12 classes in total.

2.3. Myoelectric data recording

It has been suggested that four to six surface sensors (myoelectric channels) capture adequate muscle activation information required for good classification accuracies [3,22]. Any further increase in the number of channels may not necessarily increase the classification accuracy [23,24]. We used wireless sensors to record and transmit myoelectric data at a sampling rate of 1500 Hz [25,26] with a Noraxon TeleMyo Direct Transmission System (DTS) (Noraxon USA. Inc., Scottsdale, Arizona). We used disposable, self-adhesive silver/silver chloride (Ag/AgCl) snap electrodes with two circular conductive areas of 1 cm each and an inter-electrode distance of 2 cm. It is known that a comparable accuracy can be achieved without targeting specific muscle sites [23]. Therefore, we placed all five electrodes around the circumference of the forearm symmetrically with the first electrode placed beneath the medial epicondyle of the humerus. Electrodes were placed proximally around the forearm at a location that is 1/3 of the distance between medial epicondyle of the humerus and styloid process of the ulna. Electrode placement for the data collection is shown in Fig. 2(a) and (b).

2.4. Data collection protocol

Before start of the data collection experiment, each participant was sitting comfortably in a chair with adjustable height armrests. Height of the right armrest was adjusted for each individual keeping in view his/her comfort and ease. Initial/rest position of the arm was defined as: the dorsopalmar axis pointing inside parallel to the coronal plane (palm facing to the medial side of the body with forearm parallel to the ground), elbow flexed at 90° and arm abducted 10°. Forearm was supported at two places, i.e., under the styloid process of the ulna and the distal end of the humerus (the elbow joint). A graphical user interface (GUI) based software was developed to provide visual and auditory cues to participants for guiding through the data collection process (Fig. 2(c)). The GUI provided the flexibility to control different experimental parameters including the movement duration, the number of movement repetitions in a trial and the duration of the inter-movement break. For this data collection, a single trial consisted of one repetition of

(a) (b) (c)

Fig. 2. (a) Symmetrical placement of electrodes around the forearm on anterior and posterior sides. (b) Arrangement of wireless myoelectric sensors on the forearm of a participant. (c) A snapshot of the graphical user interface (GUI) for data collection. The GUI provided both visual and auditory cues for movement initiation and termination. Duration of movement, number of repetitions of a movement in a single trial and inter-movement break duration can be specified separately for each trial by the experimenter.

each movement appearing in front of the participant in a random order. There was a short break of 4 s between consecutive movements. Once indicated by the software through visual and auditory cues, the user performed a movement for a duration of 5 s [6,12,27]. Between every two trials, there was an additional break, the duration of which was left at the discretion of the participant. A total of ten trials were recorded for each participant. Participants were instructed to start the movement from defined initial (rest) position, stay in the prescribed posture (shown in the GUI) and then return back to initial position. Participants were instructed to maintain comfortable and repeatable force levels for all movements. Furthermore, participants were asked to report any signs of pain and/or fatigue immediately to the experimenter. Before start of the data collection, all participants were given approximately 10 min of training to familiarize them with the experimental protocol and movements to be performed. Myoelectric data was collected using the Noraxon software (MyoResearch XP) and stored in the hard disk for later processing. All data processing was performed in Matlab (The MathWorks, Natick, MA).

2.5. Myoelectric signal processing

We developed a custom software in Matlab for processing of the myoelectric data, i.e., data segmentation, feature extraction and classification. Myoelectric data from a 3 s period was extracted from the raw myoelectric data (originally recorded for 5 s) from all five channels. The data was segmented to form non-overlapping analysis windows of 200 ms.

3. Heteroscedastic processes and myoelectric signal modeling

Heteroscedastic processes are characterized by a volatile nature, and are often encountered in econometrics and finance, as for instance in stock prices, which exhibit periods of large inter-day price variability followed by periods of relative stability. The ARCH and GARCH processes are used to model heteroscedastic time-series [28,29]. There are numerous generalizations of these processes, e.g., nonlinear asymmetric GARCH, integrated GARCH, exponential GARCH, quadratic GARCH, Glosten-Jagannathan-Runkle GARCH and threshold GARCH [30]. In this paper, we focus on the GARCH process to model heteroscedasticity exhibited by the myoelectric signal.

3.1. The GARCH process

Let Z_t be a sequence of *i.i.d.* random variables with zero mean and unit variance from some specified probability density function. The process Y_t is called GARCH(p, q) if

$$Y_t = \sigma_t Z_t, \qquad t \in \mathcal{Z}, \tag{1a}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{1b}$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$, $\forall\ i, j$; $\sigma_t \geq 0$ and $Z_t \sim \mathcal{N}(0, 1)$. If the parameter $p = 0$ in the second summation in (1b), the GARCH(p, q) process reduces to an ARCH(q) process [28]. In an ARCH(q) process, the conditional variance is a linear function of past sample variances only, whereas the GARCH(p, q) process takes into account lagged conditional variances as well [29]. If both $p = q = 0$, the process Y_t is a pure white noise with variance specified by the parameter α_0. Some relevant properties of the GARCH process are discussed in Appendix B, where we show that the GARCH is a white noise process.

3.2. The AR-GARCH process

AR processes are driven by white noise [31] and we have shown that the GARCH process is white noise. We can, therefore, generalize an AR process by considering a white noise of the form specified by a GARCH process. The resulting process is called the AR-GARCH, where the AR residuals ϵ_t are given by $\sigma_t Z_t$ and σ_t follows a GARCH volatility specifications in terms of historical values of ϵ_t. This formulation is a generalization of the family of AR and GARCH models to exploit features of both models [32]. For an observed signal X_t, generated by an AR(m) and a GARCH(p, q) process,

$$X_t = \sum_{k=1}^{m} \phi_k X_{t-k} + \epsilon_t, \tag{2a}$$

$$\epsilon_t = \sigma_t Z_t, \tag{2b}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{2c}$$

where α_i, β_j and Z_t are defined in Eq. (1).

3.3. Modeling heteroscedastic processes

The first step towards modeling a heteroscedastic time-series is to confirm the presence of heteroscedasticity. Once confirmed, different statistical tests can be performed to model and quantitatively measure the validity of the modeling process. We discuss four statistical tests to establish the heteroscedastic nature of the myoelectric signal and later model the same as an AR-GARCH process. These tests include the Ljung–Box Q-test, Engle's test of heteroscedasticity, the Kolmogorov–Smirnov test, and the *goodness of fit* test which are discussed in the following.

3.3.1. The Ljung–Box Q-test

The Ljung–Box Q-test is a type of portmanteau statistical test of whether any of a group of autocorrelations for a time-series are different from zero [33]. The outcome of the Ljung–Box Q-test is a parsimonious model for which the sample autocorrelation of the residuals falls below the critical value (determined by the significance level), i.e., the modeling residuals become approximately white noise. Further details of the test are given in Appendix C. We employ the Ljung–Box Q-test to find a statistically correct AR model order for the myoelectric signal.

3.3.2. Engle's test of heteroscedasticity

Robert Engle, in his seminal paper, proposed a test to establish the presence of heteroscedasticity in a given time-series [28]. According to this test, in an ARCH model setting, under the null hypothesis all ARCH parameters will be zero and under the alternate hypothesis, at least one parameter will be nonzero. Formally, we have

$$H_0 : \alpha_1 = \cdots = \alpha_p = 0,$$
$$H_1 : \alpha_i \neq 0, \text{ for some } i = 1, \ldots, p. \tag{3}$$

The procedure is to run ordinary least squares regression and find the residuals. The next step is to regress the squared residuals on a constant and p lags to test for *Lagrange multiplier* statistic TR^2 as *chi-square* (χ^2_p), where T is the sample size and R^2 is the coefficient of determination of regression. Under the null hypothesis, the asymptotic distribution of the test statistic is χ^2_p with p degrees of freedom. We employ the Engle's test to confirm that the AR residuals are heteroscedastic.

3.3.3. The Kolmogorov–Smirnov test

The two sample Kolmogorov–Smirnov (K–S) test is a non-parametric test for the equality of continuous, one-dimensional probability distributions that can be used to compare two samples. We use the K–S test to quantitatively measure how close the AR residual distribution is to that of a GARCH process. Further details about the test are presented in Appendix D.

3.3.4. Goodness of fit

The *goodness of fit* of a statistical model describes how well the model fits the observations. In our setting, we need to quantitatively measure how well the AR-GARCH model fits the myoelectric signal. There are various tests available in the literature which can be used to find the *goodness of fit* of a proposed statistical model [34]. We used the Akaike information criteria (AIC) [35] and the Bayesian information criteria (BIC) [36]. Lower AIC/BIC values correspond to a better fit of the model to the observed data.

4. Results

We present results in two separate parts. In the first part, we analyzed the myoelectric signal for heteroscedasticity and modeled it as an AR-GARCH process. Later, we used the AR-GARCH model parameters as a new feature set and performed classification of the myoelectric signal using two different classification algorithms.

4.1. Myoelectric signal modeling and analysis

We modeled the myoelectric signal as an AR process with the condition that the AR modeling residuals were white noise. Later, we analyzed the AR modeling residuals for heteroscedasticity and modeled the residuals as a GARCH process. Finally, we performed a *goodness of fit* analysis of the proposed model.

4.1.1. Myoelectric signal as an AR process

In order to find a statistically correct AR model, we increased the AR model order till the modeling residuals became white noise. Specifically, for each AR model, the residuals formed an input to the Ljung–Box Q-test which analyzed residuals. In addition, we also used squared values of the residuals as an input to the Ljung–Box Q-test. The later input was used to highlight the heteroscedastic nature of the myoelectric signal. We used a 95% significance level for the test, with the AR model order ranging from 5 to 15. An analysis window of 200 ms was used in this test.

In Fig. 3, it is evident that for the AR(9) model, the residuals are approximately white noise, while the Q-statistic for the squared AR residuals does not decrease with increasing model order, which is

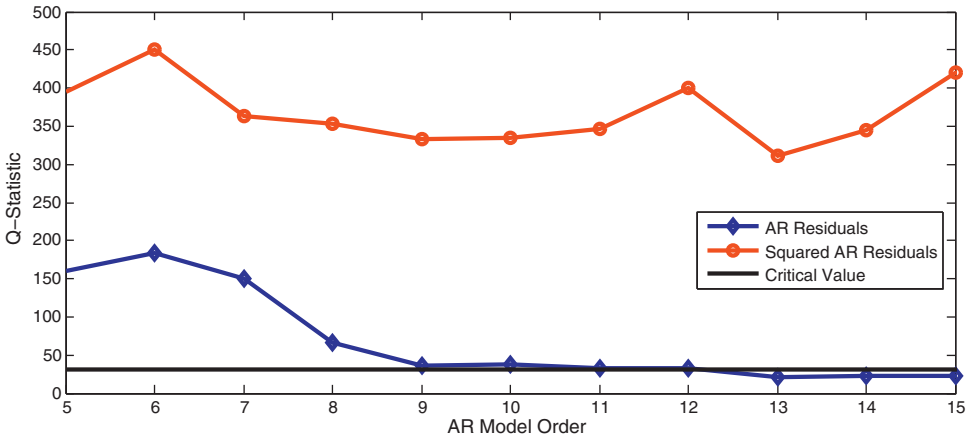

Fig. 3. Ljung–Box Q-test results for the myoelectric signal. The AR modeling residuals are approximately white noise for AR(9), while for the squared residuals, the Q-statistic does not decrease and stays above the critical value for all AR models. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)

Table 1
Heteroscedasticity in the myoelectric signal for different analysis windows and AR models. Myoelectric data was recorded at a sampling frequency of 1500 Hz. Critical value for all tests was 18.71.

Window size	Samples	Statistic value	Result
AR(5)			
50 ms	75	11.55	Homoscedastic
100 ms	150	29.44	Heteroscedastic
150 ms	225	48.37	Heteroscedastic
200 ms	300	44.86	Heteroscedastic
250 ms	375	46.27	Heteroscedastic
300 ms	450	48.84	Heteroscedastic
AR(9)			
50 ms	75	10.24	Homoscedastic
100 ms	150	33.08	Heteroscedastic
150 ms	225	55.53	Heteroscedastic
200 ms	300	40.79	Heteroscedastic
250 ms	375	50.03	Heteroscedastic
300 ms	450	54.66	Heteroscedastic

a typical characteristic of a heteroscedastic time-series (in accordance with the lower right figure in Fig. B.9).

4.1.2. Heteroscedasticity in the myoelectric signal

Once a statistically correct AR model was found (i.e., the modeling residuals were white noise), we used the Engle's test to establish whether the residuals exhibited heteroscedasticity or otherwise. For this test, we used analysis windows of various sizes and two different AR models, i.e., AR(5) and AR(9). In Table 1, we present results of the heteroscedastic testing. We varied the size of analysis windows from 50 ms (75 samples) to 300 ms (450 samples). For both AR models and all analysis windows of 100 ms or above, the null hypothesis was rejected, i.e., the myoelectric signal was found to exhibit heteroscedasticity.

Further, we used a two sample K–S test to quantify whether the AR modeling residuals follow the same distribution as that of a GARCH process. For this purpose, one sample for the K–S test consisted of the AR residuals and the other sample was the GARCH process generated with parameters estimated using the same AR residuals. The test was performed on various myoelectric signals selected at random from experimental data and modeled with an AR(5) or AR(9) process. Invariably, in all cases, the test was unable to reject the null hypothesis indicating that the GARCH process and the residual samples were following the same probability distribution function. In Fig. 4, we present the cumulative distribution function (CDF) and histogram plots for the AR(5) residuals and the GARCH(1,1) process generated using α_0, α_1 and β_1 parameters estimated from the same AR modeling residuals and Gaussian innovations for Z_t (Eq. (1)).

4.1.3. Goodness of fit

Finally, to establish whether the proposed AR-GARCH modeling scheme actually fits the data well, we perform *goodness of fit* testing using the AIC and BIC. We used different number of models, i.e., AR(m), AR(m)-GARCH(1,1) and AR(m)-GARCH(2,2) with $m = 2, \ldots 12$. The results from the *goodness of fit* testing are presented in Fig. 5.

We have shown that the myoelectric signal can be modeled as an AR process, however, a statistically correct model order may be higher, i.e., AR(9), than what is generally used in the literature, i.e., AR(4) to AR(6). Further, we found that the AR residuals exhibited heteroscedasticity and can be modeled as a GARCH process. We also established that the models AR(m)-GARCH(1,1) and AR(m)-GARCH(2,2) with $m > 8$ best fitted the myoelectric signal. Similarly, the Ljung–Box Q-test resulted in an AR(9) model, therefore we selected an AR(9) model for further analysis. Also, for the GARCH part, we selected GARCH(1,1) being computationally less expensive to estimate. A summary of presented statistical tests and their outcome is shown in Table 2. Based on the presented analysis, we selected the AR(9)-GARCH(1,1) model coefficients as a new feature set for myoelectric signal classification.

4.2. Myoelectric feature extraction and classification

For the purpose of feature extraction from the myoelectric signal, we used an analysis window of size 200 ms. After forming analysis windows, the myoelectric signals were modeled as an

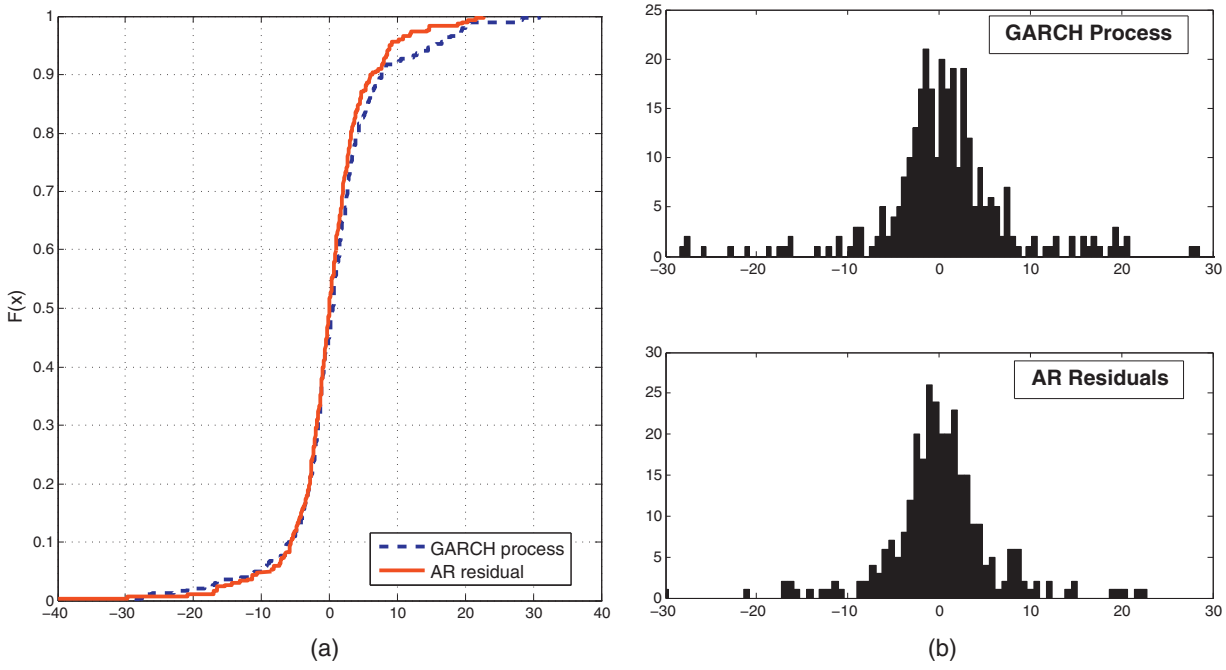

Fig. 4. CDFs and histogram plots for the GARCH(1,1) process and the AR(5) residuals. The GARCH(1,1) process was generated using the parameters estimated from the same AR residuals and Gaussian innovations.

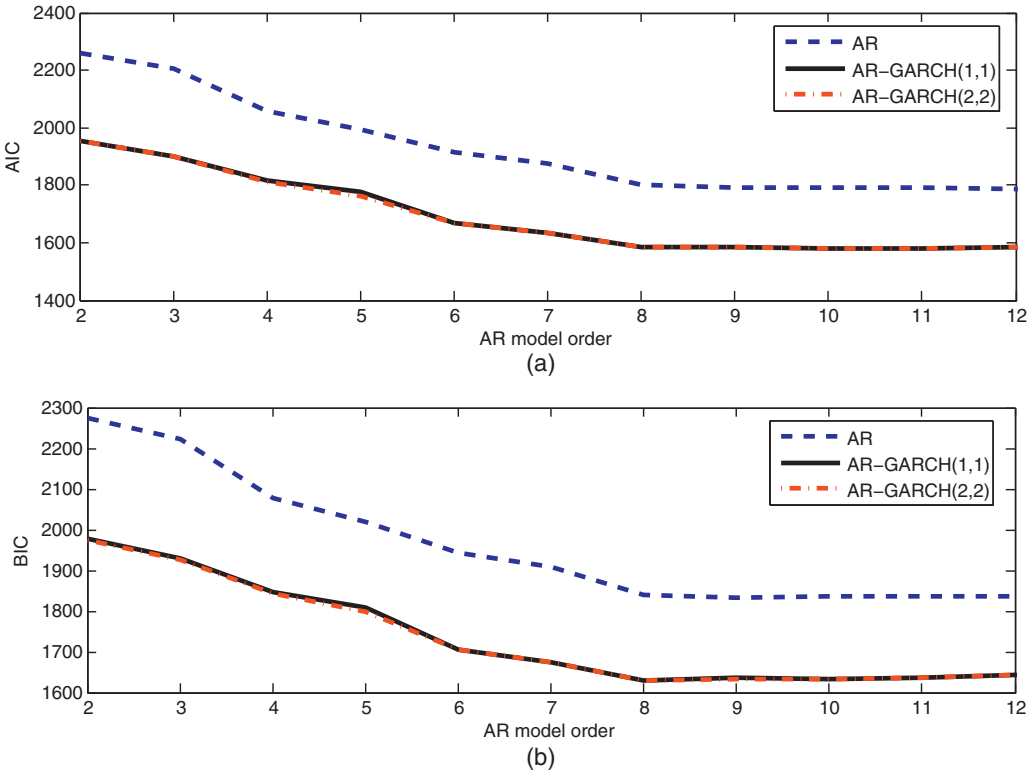

Fig. 5. Results of *goodness of fit* using the AIC and the BIC for the AR(m), AR(m)-GARCH(1,1) and AR(m)-GARCH(2,2) with $m = 2, \cdots, 12$ models. There is a significant difference between the AR(m) and the AR(m)-GARCH models. However, the AR(m)-GARCH(1,1) and AR(m)-GARCH(2,2) are close to each other indicating a diminishing effect of increasing the GARCH model order from (1,1) to (2,2).

Table 2
A summary of statistical tests to model the myoelectric signal as an AR-GARCH process. All four tests were performed sequentially to find a parsimonious model for the myoelectric signal.

Statistical test	Outcome
Ljung–Box Q-test	The test ascertains a parsimonious model order for the myoelectric signal once modeled as an AR process. We found that for the analyzed myoelectric signal, an AR(9) model resulted in white noise residual.
Engle's test	The test is used on the AR modeling residuals to establish whether the residuals exhibited heteroscedasticity or otherwise. We found that the AR residuals of the myoelectric signal actually exhibits heteroscedasticity and we may model the residuals as a GARCH process.
K–S test	The test determines whether the AR residuals follow the same probability distribution as that of a GARCH process. The test confirmed our hypothesis that both AR residuals and the GARCH(1,1) process actually follow the same distribution.
Goodness of fit	AIC and BIC were used to measure the *goodness of fit* for the myoelectric signal when modeled as an AR or AR-GARCH process. We found that for the analyzed myoelectric signal, an AR(9)-GARCH(1,1) model performed better than others.

AR(m) process and $m+1$ coefficients were estimated (Eq. (2a)). Next, the AR residuals ϵ_t were modeled as a GARCH(p, q) process and $p+q+1$ coefficients were estimated (Eq. (2c)). Two additional coefficients are the AR and the GARCH model constants. The proposed AR(9)-GARCH(1,1) resulted into a total of 13 coefficients. For the purpose of comparison, we consider three different feature sets, i.e., AR(5), AR(9), and AR(9)-GARCH(1,1). A summary of the selected feature sets is given in Table 3.

To evaluate the performance of the selected three feature sets, we employed two widely used classification algorithms, i.e., the linear discriminant analysis (LDA) and the artificial neural network (ANN) [2,5]. The LDA was used in a five-fold cross-validation setting, where all the data was randomly partitioned into five approximately equal size bins. Out of five data bins, a single data bin was retained as the validation data for testing, and the remaining four data bins were used for training. The cross-validation process was repeated 10 times. The classification errors were averaged to get the final classification error. Similarly for the ANN classifier, we used a feedforward network with one input layer, one hidden layer with six neurons and an output layer. Back-propagation training algorithms, Levenberg–Marquardt was used to train the neural network [37]. The ANN was trained ten times and an average of the classification error was recorded. Classification errors for individual participants and overall averages are presented in Fig. 6.

Table 3
Feature sets used in the classification.

Feature set	Reason
AR(5)	Prevalent use in the literature [5].
AR(9)	Resulting model from the Ljung–Box Q-test having white noise residuals (Fig. 3).
AR(9)-GARCH(1,1)	A model with white noise residuals that were further modeled as a GARCH(1,1) process for heteroscedasticity.

We performed the multivariate analysis of variance (MANOVA) with the participants and the feature sets as independent variables and classification errors of the LDA and the ANN as dependent variables. We observed a significant effect of feature sets on the classification accuracy for both algorithms ($p < .01$). A pairwise comparison revealed that the means of the LDA and the ANN

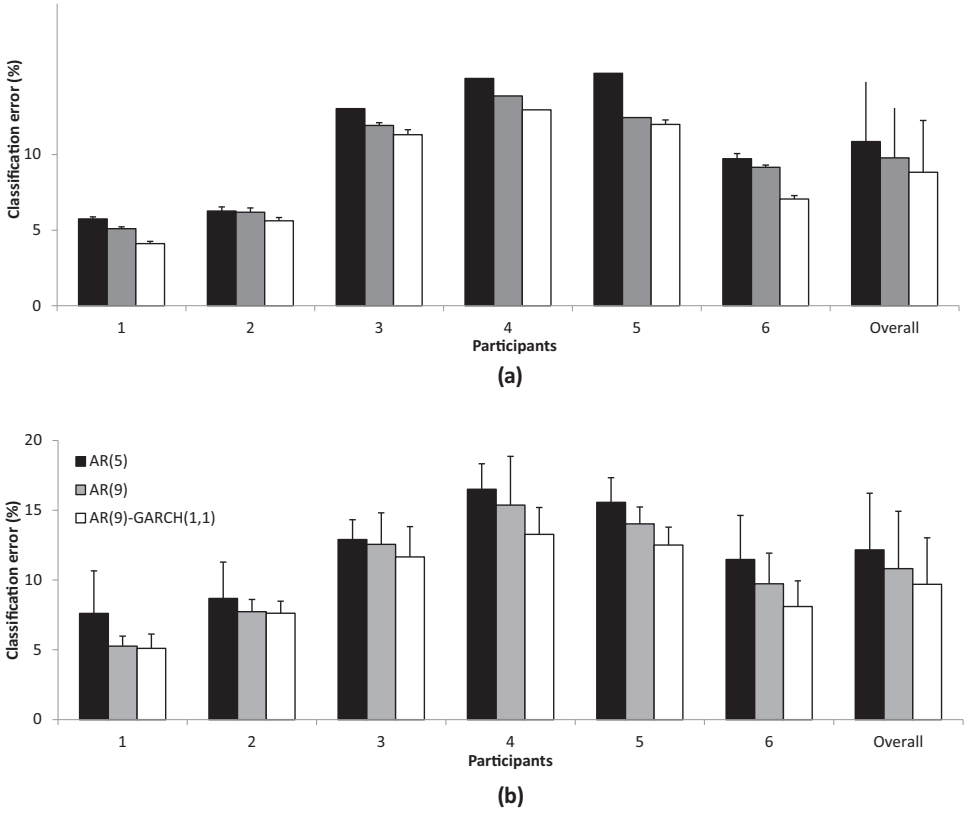

Fig. 6. Average classification errors for all participants and overall classification error using three sets of features AR(5), AR(9), and AR(9)-GARCH(1,1).

classification errors were significantly different for all three feature sets ($p < .01$ for all pairwise comparisons). In other words, the classification accuracy was significantly affected by the choice of the feature set, i.e., AR(5), AR(9) or AR(9)-GARCH(1,1) for both the LDA and the ANN algorithms (Fig. 7).

To test and compare the reliability of myoelectric classification for all three feature sets, we used two standard metrics, i.e., the sensitivity and the specificity defined as,

$$\text{Sensitivity} = \frac{\text{Number of recognized true movements}}{\text{Number of true movements}}, \quad (4)$$

$$\text{Specificity} = \frac{\text{Number of rejected false movements}}{\text{Number of false movements}}. \quad (5)$$

The average sensitivity and specificity values over all movements are shown in Table 4. It is evident that the newly proposed feature set, i.e. AR(9)-GARCH(1,1) has higher average sensitivity and specificity values over all movements, i.e., 93.03% and 91.36% respectively. The higher values show that the proposed feature set is more reliable compared to the classical AR feature sets. Confusion matrices for all three feature sets using the LDA classifier are shown in Fig. 8.

Table 4
Average sensitivity and specificity values for the LDA algorithm across all movements.

Feature set	Sensitivity	Specificity
AR(5)	90.80%	89.14%
AR(9)	91.83%	90.17%
AR(9)-GARCH(1,1)	93.03%	91.36%

5. Discussion

We analyzed the myoelectric signal recorded during performance of different movements and found that the signal exhibited heteroscedasticity. A detailed investigation using statistical tests including the Ljung–Box Q-test, the Engle's test, the

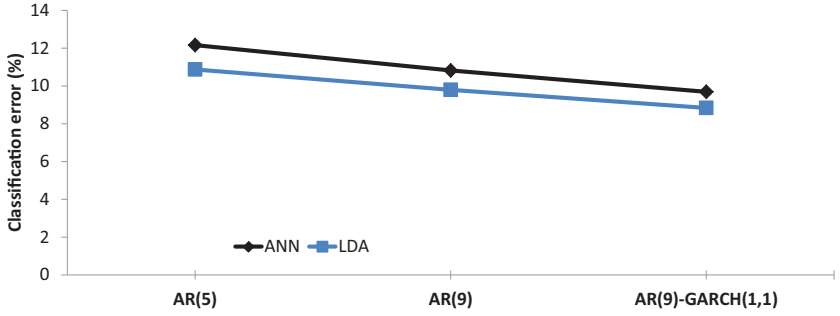

Fig. 7. Average classification errors using three feature sets, i.e., AR(5), AR(9), and AR(9)-GARCH(1,1).

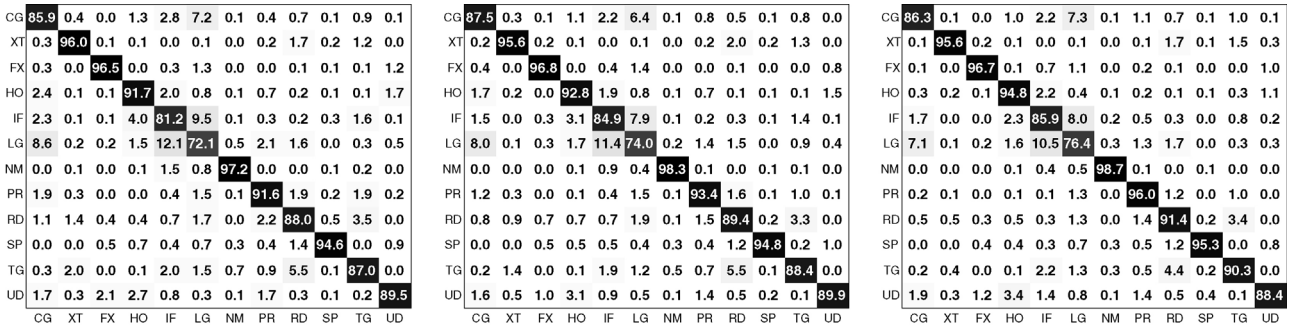

Fig. 8. Confusion matrices for the AR(5), AR(9) and AR(9)-GARCH(1,1), where CG:Close Grip, XT:Extension, FX:Flexion, HO:Hand Open, IF:Index Finger Point, LG:Lateral Grasp, NM:No Movement, PR:Pronation, RD:Radial Deviation, SP:Supination, TG:Tripod Grip, UD:Ulnar Deviation.

Kolmogorov–Smirnov test and *goodness of fit* revealed that an AR(9)-GARCH(1,1) best fitted the myoelectric data. The prevalent use of AR(4)–AR(6) [5] in the literature does not capture all the information present in the myoelectric signal. Using the Ljung–Box *Q*-test, we found that an AR(9) model captures enough information from the myoelectric signal that the residuals become white noise. However, the residuals for AR(5) as well as AR(9) exhibited heteroscedastic behavior as evident from Table 1 and Fig. 3 (red line with circular markers). It is evident that increasing the model order does not have a significant effect on the heteroscedasticity of the myoelectric signal. Therefore, we assume that it is imperative to consider the GARCH modeling of the AR residuals to capture an increased amount of information from the myoelectric signal.

The window size controls the amount of raw myoelectric information available to the feature extraction routine; and therefore directly affects the classifier performance [1,38]. An analysis window of size ranging from 150 ms to 250 ms is considered to give good classification accuracy [38]. Our intensive investigation and experiments with myoelectric signals revealed that an analysis window of 200 ms provided maximum classification accuracy. As the information contents in the myoelectric signal change due an increase/decrease in the analysis window size, we assume that the correct AR model order will depend on analysis window size. Therefore, we recommend that for a different analysis window size, the Ljung–Box *Q*-test should be performed again to ascertain a correct AR model order. Once an AR model has been ascertained then the residuals should be modeled as a GARCH process. Furthermore to determine correct GARCH model order, we recommend using some information criteria, e.g., the AIC or the BIC (Fig. 5).

We also used the K–S test to confirm whether the AR residuals and the GARCH process exhibited similar statistical properties. The K–S test as well as a visual analysis (Fig. 4) of both data sequences revealed that the both may have same probability distribution function. Therefore, our assumption for modeling the AR residuals using the GARCH process is supported.

Apart from features, classification algorithms are also an important part of myoelectric signal classification schemes. Details of different classifies can be found in Refs. [2,3,5] and references therein. This study was primarily focused on AR features and heteroscedastic modeling of the myoelectric signal. Therefore, we focused on two common classifiers [2,3,5].

6. Conclusion

We focused on the modeling of the myoelectric signal used for classification of different hand and wrist movements in myoelectric prostheses. We discovered that once the myoelectric signal is modeled as an AR process, the modeling residuals exhibit heteroscedasticity, i.e., the average size of residuals is not constant. Subsequently, we proposed the GARCH process to model such residuals. We performed multiple statistical tests to establish the heteroscedasticity of the myoelectric signal and showed the applicability of the proposed AR-GARCH model. The coefficients of the AR-GARCH model formed the new feature set for the pattern classification. Experimental results, conducted on several participants, show that the newly proposed feature set significantly outperformed ($p < .01$) the conventionally adopted AR feature set in classifying 12 hand and wrist movements.

Appendix A. Concept of heteroscedasticity

In order to illustrate the concept of heteroscedasticity, let us consider a simple linear regression model between two variables x and y,

$$y = \alpha + \beta x + \epsilon, \tag{A.1}$$

where ϵ represents the regression error. We call *homoscedasticity* the assumption that the expected size of the error is constant. We call *heteroscedasticity* the assumption that the expected size of the error term is not constant [39]. The assumption of homoscedasticity is standard in regression theory because of its mathematical convenience. However, in many applications, this assumption may be unreasonable. A classic example of heteroscedasticity is that of income versus expenditure. Those with higher income will display a greater variability in consumption than lower-income individuals who tend to spend a rather constant amount. The concept of heteroscedasticity generalizes to many other linear and non-linear models including the AR model. The presence of heteroscedasticity in a time-series can invalidate statistical tests that assume that the model residual variances are uncorrelated. It is possible to specify a stochastic process for the residual errors and predict the average size of the error terms [39] using the Autoregressive Conditional Heteroscedasticity (ARCH) [28] or the Generalized ARCH (GARCH) models [29]. The GARCH process is a generalization of the ARCH process, and can model the heteroscedasticity more parsimoniously.

Appendix B. Properties of the GARCH process

The expected value ($E[.]$) of the GARCH (p, q) process is given by

$$E[Y_t] = E[\sigma_t Z_t] = E[\sigma_t]E[Z_t] = 0. \tag{B.1}$$

Similarly, the variance ($V[.]$) of the GARCH (p, q) process is given by

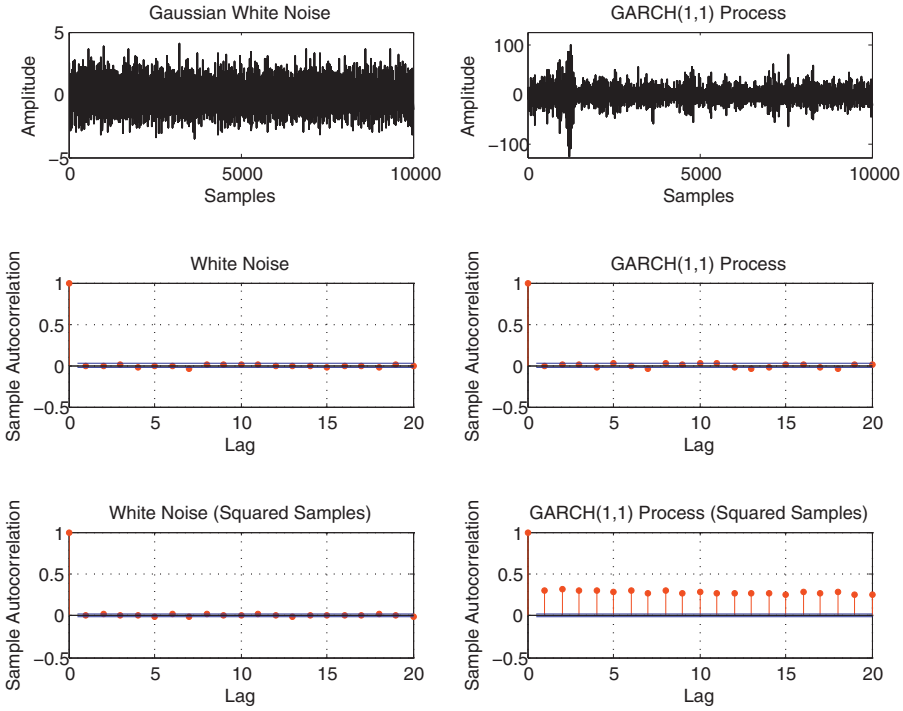

Figure B.9. A comparison of the Gaussian white noise and the GARCH(1,1) process. The first column presents Gaussian white noise and the second column presents GARCH(1,1) process. Top row shows both processes. In the middle row, we present the samples autocorrelation values for 20 lags for both processes. The last row presents the sample autocorrelations of the squared processes. We can see that the GARCH process is a white noise just like the Gaussian white noise, but with a unique characteristic, i.e., there is a significant autocorrelation in its squared sample values as is evident in the lower right figure. The Gaussian noise was generated using standard normal distribution, and the GARCH process was generated using $Z_t \sim \mathcal{N}(0, 1)$, $\alpha_0 = 2$, $\alpha_1 = 0.09$ and $\beta_1 = 0.9$.

$$\begin{aligned} V[Y_t] &= E[Y_t^2] - [E(Y_t)]^2 = E[Y_t^2] = E[\sigma_t^2], \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i E[Y_t^2] + \sum_{j=1}^{p} \beta_j E[\sigma_t^2], \\ &= \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}. \end{aligned} \tag{B.2}$$

From Eqs. (B.1) and (B.2), it is clear that the GARCH(p, q) is a zero mean process with a constant variance, which is specified by the parameters $\{\alpha_i\}_{i=0}^{q}$ and $\{\beta_j\}_{j=1}^{p}$. The GARCH process is also statistically uncorrelated, i.e.,

$$E[Y_t Y_s] = E[\sigma_t Z_t \sigma_s Z_s] = E[\sigma_t Z_t \sigma_s] E[Z_s] = 0. \tag{B.3}$$

We show next that the conditional variance of the GARCH process is time-varying, reflecting the volatility of the stochastic signal. Let $\mathfrak{F}_t = \{Y_s : s \leq t\}$ represent the history of the process up to time t. We have

$$E[Y_t|\mathfrak{F}_{t-1}] = E[\sigma_t Z_t|\mathfrak{F}_{t-1}] = \sigma_t E[Z_t|\mathfrak{F}_{t-1}] = \sigma_t E[Z_t] = 0 \tag{B.4}$$

$$V[Y_t|\mathfrak{F}_{t-1}] = E[\sigma_t^2 Z_t^2|\mathfrak{F}_{t-1}] = \sigma_t^2 V[Z_t] = \sigma_t^2 \tag{B.5}$$

From (B.4) and (B.5), we see that the conditional expected value of the GARCH process is also zero but the conditional variance is equal to σ_t^2, which is a time-varying quantity. In summary, we have shown that the GARCH(p, q) process is a white noise process with time-varying conditional variance.

To understand the difference between a white noise process and the GARCH process, in Fig. B.9, we generated a Gaussian white noise, w_t with zero-mean and unit variance, and a GARCH process, Y_t, using Gaussian innovations. A total of 10,000 realizations for each process were generated. The sample autocorrelations of the two precesses, w_t, Y_t, and the squared processes, w_t^2, Y_t^2, were computed for 20 lag values. It is evident from Fig. B.9 that the squared sample autocorrelation of the GARCH process is much higher than the defined significance level. Evidently, the GARCH process is a white noise with statistically significant autocorrelation in its squared sample values.

Appendix C. The Ljung–Box Q-test

The Ljung–Box Q-test assesses the null hypothesis that the series of residuals exhibits no autocorrelation for a fixed number of lags m versus the alternative hypothesis that the autocorrelation sequence is non-zero. Under the null hypothesis the asymptotic distribution of the Q-statistic is χ_m^2 with m degrees of freedom. The modified Q-statistic is defined as,

$$Q(r) = n(n+2) \sum_{k=1}^{m} \frac{r_k^2}{n-k}, \tag{C.1}$$

$$r_k = \sum_{t=k+1}^{n} \frac{\epsilon_t \epsilon_{t-k}}{\sum_{t=1}^{n} \epsilon_t^2},$$

where r_k is the autocorrelation estimate of the residuals at lag k, m is an appropriate number of lags of the sample autocorrelation, n is the sample size, and ϵ represents the modeling residuals.

Appendix D. The Kolmogorov–Smirnov test

The Kolmogorov–Smirnov (K–S) statistic quantifies a distance between the empirical distribution functions of two samples. The null distribution of this statistic is calculated under the null

hypothesis that the samples are drawn from the same distribution [40]. For the two sample test, the K–S statistic is given by:

$$D_{m,n} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|, \tag{D.1}$$

where $F_{1,n}$ and $F_{2,m}$ are the empirical cumulative distribution functions of the first and second sample respectively. The empirical cumulative distribution function is given by

$$F_k(t) = \frac{1}{k}\sum_{i=1}^{k} \mathbf{1}\{x_i < t\}, \tag{D.2}$$

where $\mathbf{1}\{.\}$ is the indicator function and is equal to 1 for $x_i < t$ and 0 otherwise. The null hypothesis is rejected at level α if

$$\sqrt{\frac{n+m}{nm}} D_{n,m} > K_\alpha, \tag{D.3}$$

where K_α is the critical value calculated from the defined significance level α [41].

References

[1] K. Englehart, B. Hudgins, A robust, real-time control scheme for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 50 (2003) 848–854.
[2] S. Micera, J. Carpaneto, S. Raspopovic, Control of hand prostheses using peripheral information, IEEE Rev. Biomed. Eng. 3 (2010) 48–68.
[3] E. Scheme, K. Englehart, Electromyogram pattern recognition for control of powered upper-limb prostheses: state of the art and challenges for clinical use, J. Rehabil. Res. Dev. 48 (2011) 643.
[4] N. Jiang, K.B. Englehart, P.A. Parker, Extracting simultaneous and proportional neural control information for multiple-DOF prostheses from the surface electromyographic signal, IEEE Trans. Biomed. Eng. 56 (2009) 1070–1080.
[5] B. Peerdeman, D. Boere, H. Witteveen, R.H. in 't Veld, H. Hermens, S. Stramigioli, H. Rietman, P. Veltink, S. Misra, Myoelectric forearm prostheses: state of the art from a user-centered perspective, J. Rehabil. Res. Dev. 48 (2011) 719–737.
[6] M.A. Oskoei, H. Hu, Support vector machine-based classification scheme for myoelectric control applied to upper limb, IEEE Trans. Biomed. Eng. 55 (2008) 1956–1965.
[7] G. Naik, D. Kumar, Jayadeva, Twin SVM for gesture classification using the surface electromyogram, IEEE Trans. Inf. Technol. Biomed. 14 (2010) 301–308.
[8] X. Chen, D. Zhang, X. Zhu, Application of a self-enhancing classification method to electromyography pattern recognition for multifunctional prosthesis control, J. NeuroEng. Rehabil. 10 (2013) 44.
[9] G.R. Naik, D.K. Kumar, M. Palaniswami, Surface EMG based hand gesture identification using semi blind ICA: validation of ICA matrix analysis, Electromyogr. Clin. Neurophysiol. 48 (2008) 169–180.
[10] G.R. Naik, D.K. Kumar, Identification of hand and finger movements using multi run ICA of surface electromyogram, J. Med. Syst. 36 (2012) 841–851.
[11] P.A. Parker, R.N. Scott, Myoelectric control of prostheses, Crit. Rev. Biomed. Eng. 13 (1986) 283.
[12] D. Tkach, H. Huang, T. Kuiken, Study of stability of time-domain features for electromyographic pattern recognition, J. NeuroEng. Rehabil. 7 (2010) 21.
[13] A. Young, L. Hargrove, T. Kuiken, Improving myoelectric pattern recognition robustness to electrode shift by changing interelectrode distance and electrode configuration, IEEE Trans. Biomed. Eng. 59 (2012) 645–652.
[14] G. Rasool, N. Bouaynaya, K. Iqbal, Muscle activity detection from myoelectric signals based on the AR-GARCH model, in: 2012 IEEE Statistical Signal Processing Workshop (SSP), 2012, pp. 420–423.
[15] S. Mihandoost, M. Amirani, B. Varghahan, A new approach for feature extraction of EEG signal using GARCH variance series, in: 2011 5th International Conference on Application of Information and Communication Technologies (AICT), 2011, pp. 1–5.
[16] K.F.K. Wong, A. Galka, O. Yamashita, T. Ozaki, Modelling non-stationary variance in EEG time series by state space GARCH model, Comput. Biol. Med. 36 (2006) 1327–1335.
[17] A. Galka, O. Yamashita, T. Ozaki, GARCH modelling of covariance in dynamical estimation of inverse solutions, Phys. Lett. A 333 (2004) 261–268.
[18] M. Amirmazlaghani, H. Amindavar, EMG signal denoising via Bayesian wavelet shrinkage based on GARCH modeling, in: Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Washington, DC, 2009, pp. 469–472.
[19] H. Hermens, B. Freriks, R. Merletti, D. Stegeman, J. Blok, G. Rau, C. Disselhorst-Klug, G. Hägg, European Recommendations for Surface Electromyography, Roessingh Research and Development, The Netherlands, 1999.
[20] E.F. Delagi, M.D. Iazzetti, John, M.D. Perotto, O. Aldo, M.D. Morrison, Daniel, Anatomical Guide for the Electromyographer: The Limbs and Trunk, fifth ed., Charles C Thomas Pub Ltd, Springfield, Illinois, USA, 2011.
[21] N. Jiang, J.L. Vest-Nielsen, S. Muceli, D. Farina, EMG-based simultaneous and proportional estimation of wrist/hand dynamics in uni-lateral trans-radial amputees, J. NeuroEng. Rehabil. 9 (2012) 42.
[22] G.R. Naik, D.K. Kumar, M. Palaniswami, Signal processing evaluation of myoelectric sensor placement in low-level gestures: sensitivity analysis using independent component analysis, Expert Syst. 31 (2014) 91–99.
[23] T. Farrell, R. Weir, A comparison of the effects of electrode implantation and targeting on pattern classification accuracy for prosthesis control, IEEE Trans. Biomed. Eng. 55 (2008) 2198–2211.
[24] L. Hargrove, K. Englehart, B. Hudgins, A comparison of surface and intramuscular myoelectric signal classification, IEEE Trans. Biomed. Eng. 54 (2007) 847–853.
[25] R. Merletti, P. Di Torino, Standards for reporting EMG data, J. Electromyogr. Kinesiol. 9 (1999) 3–4.
[26] G.L. Soderberg, L.M. Knutson, A guide for use and interpretation of kinesiologic electromyographic data, Phys. Ther. 80 (2000) 485–498.
[27] P. Zhou, M.M. Lowery, K.B. Englehart, H. Huang, G. Li, L. Hargrove, J.P. Dewald, T.A. Kuiken, Decoding a new neural-machine interface for control of artificial limbs, J. Neurophysiol. 98 (2007) 2974–2982.
[28] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, Econometrica 50 (1982) 987–1007.
[29] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, J. Econometr. 31 (1986) 307–327.
[30] T. Bollerslev, Glossary to ARCH (GARCH), CREATES Research Paper 49 (2008).
[31] M.H. Hayes, Statistical Digital Signal Processing and Modeling, first ed., Wiley, 1996.
[32] S. Ling, M. McAleer, On adaptive estimation in nonstationary ARMA models with GARCH errors, Ann. Stat. 31 (2003) 642–674.
[33] G.M. Ljung, G.E.P. Box, On a measure of lack of fit in time series models, Biometrika 65 (1978) 297–303.
[34] W.K. Li, Diagnostic Checks in Time Series, first ed., Chapman and Hall/CRC, 2003.
[35] H. Akaike, A new look at the statistical model identification, IEEE Trans. Autom. Control 19 (1974) 716–723.
[36] G. Schwarz, Estimating the dimension of a model, Ann. Stat. 6 (1978) 461–464.
[37] H. Yu, B. Wilamowski, Levenberg–Marquardt training, Ind. Electron. Handb. 5 (2011).
[38] L.H. Smith, L.J. Hargrove, B.A. Lock, T.A. Kuiken, Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay, IEEE Trans. Neural Syst. Rehabil. Eng. 19 (2011) 186–192.
[39] R.F. Engle, S.M. Focardi, F.J. Fabozzi, ARCH/GARCH Models in Applied Financial Econometrics, in: Handbook of Finance, John Wiley & Sons, Inc, 2008.
[40] R.H.C. Lopes, I. Reid, P.R. Hobson, The two-dimensional Kolmogorov–Smirnov test, in: XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research, Amsterdam, the Netherlands, 2007.
[41] M. Hollander, D.A. Wolfe, Nonparametric Statistical Methods, second ed., Wiley-Interscience, 1999.