

Does the variance of surface EMG signals during isometric contractions follow an inverse gamma distribution?

Akira Furui¹, *Member, IEEE*, and Toshio Tsuji¹, *Member, IEEE*

Abstract—In this paper, the validity of the stochastic model-based variance distribution of surface electromyogram (EMG) signals during isometric contraction is investigated. In the model, the EMG variance is considered as a random variable following an inverse gamma distribution, thereby allowing the representation of variations in the variance. This inverse gamma-based model for the EMG variance is experimentally validated through comparison with the empirical distribution of variances. The difference between the model distribution and the empirical distribution is quantified using the Kullback–Leibler divergence. Additionally, regression analysis is conducted between the model parameters and the statistics calculated from the empirical distribution of EMG variances. Experimental results showed that the inverse gamma-based model is potentially suitable and that its parameters can be used to evaluate the stochastic properties of the EMG variance.

I. INTRODUCTION

The surface electromyogram (EMG) is electrical activity generated by muscle contraction and measured from the skin surface. The property of EMG signals reflects the recruitment and firing mechanism of a motor unit, which is a group of muscle fibers innervated by a single motor neuron.

The relationship between the amplitude features of EMG signals (e.g., variance and integrated value) and muscle activity has long been studied. The EMG amplitude is strongly related to muscle force [1], [2]. Furthermore, muscle activity and its underlying motor unit activity involve uncertainty, and it has been suggested that this affects the properties of EMG signals [3]. Therefore, if such uncertainty included in EMG signals could be theoretically evaluated, it would lead not only to a deeper understanding of the neuromuscular system but also applications to practical fields, such as prosthesis control and diagnosis support.

To represent such uncertainty, the authors previously proposed a stochastic model of EMG signals based on a scale mixture distribution with the assumption that the EMG variance is a random variable [4]. Through a series of experiments, it was shown that uncertainty in the EMG variance may relate to the features of motor unit activity potentially included in the EMG signals. This model assumes that the EMG variance is a random variable that follows an inverse gamma distribution; however, no investigation has been performed on the validity of this assumption.

In this paper, the validity of the inverse gamma assumption for the EMG variance in the scale mixture model [4] that

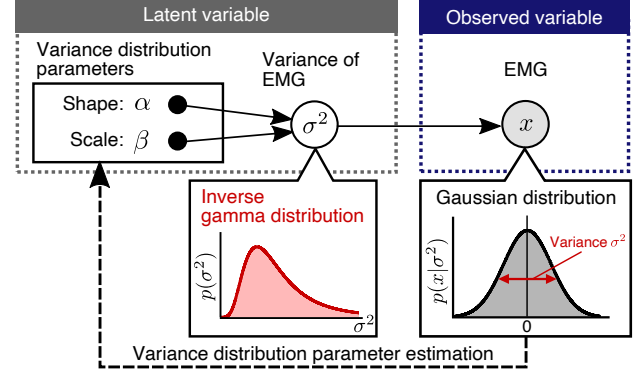


Fig. 1. Overview of the scale mixture model of EMG signals [4]. The graphical model represents the stochastic relationship between the EMG signal and its variance. The white nodes are random variables and the black nodes are parameters to be estimated. In the model, the EMG signal x is an observed variable and the EMG variance σ^2 is a latent variable.

the authors proposed is examined. The inverse gamma-based variance distribution is evaluated through a comparison with the empirical distribution of variances. Additionally, regression analysis is conducted between the model parameters and the statistics calculated from the empirical distribution, and their relationship is discussed.

II. METHODS

A. Scale Mixture Model with Variance Distribution

1) *Model structure*: The scale mixture model [4] describes the stochastic relationship between the EMG signal x and its variance σ^2 . Fig. 1 shows the structure of the scale mixture model. Both x and σ^2 are regarded as random variables, thereby allowing the representation of variations in the EMG variance. In the model, the variance σ^2 is regarded as a latent variable because it cannot be directly observed.

In the model, the distribution of the EMG signal x can be expressed in the following form:

$$p(x) = \int p(x|\sigma^2)p(\sigma^2)d\sigma^2. \quad (1)$$

Here, the conditional distribution of an EMG signal x given σ^2 is expressed as a Gaussian distribution with a mean of zero and a variance of σ^2 :

$$p(x|\sigma^2) = \mathcal{N}(x|0, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2}{2\sigma^2}\right]. \quad (2)$$

The EMG variance σ^2 is assumed to follow an inverse gamma distribution [4], [5]:

$$p(\sigma^2) = \text{IG}(\sigma^2|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left[-\frac{\beta}{\sigma^2}\right], \quad (3)$$

This work was partially supported by Grant-in-Aid for JSPS Research Fellow 18J22370.

¹A. Furui and T. Tsuji are with the Graduate School of Engineering, Hiroshima University, Higashihiroshima 739-8527, Japan (e-mail: akirafurui@hiroshima-u.ac.jp).

where $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}^+$ determine the inverse gamma distribution and are referred to as the shape and scale, respectively. The inverse gamma distribution is defined in the positive region and is known as a conjugate distribution for the variance of a Gaussian distribution. From the above, (1) can be expanded as

$$\begin{aligned} p(x) &= \int \mathcal{N}(x|0, \sigma^2) \text{IG}(\sigma^2|\alpha, \beta) d\sigma^2 \\ &= \frac{\beta^\alpha \Gamma(\alpha + \frac{1}{2})}{\sqrt{2\pi} \Gamma(\alpha) (\beta + \frac{x^2}{2})^{\alpha + \frac{1}{2}}}. \end{aligned} \quad (4)$$

From this equation, the scale mixture model is characterized by the variance distribution parameters α and β . Additionally, the model is obtained by an infinite mixture of Gaussians having different variances that are generated by the inverse gamma distribution. The scale mixture model is capable of representing changes in the distribution shape of EMG signals associated with muscle activity by considering the stochastic fluctuation of EMG variance.

2) *Parameter estimation*: Given N samples of EMG signals $X = \{x_n\}_{n=1}^N$, the latent variance distribution parameters, α and β , can be estimated by maximizing the log-likelihood $\ln p(X) = \sum_{n=1}^N \ln p(x_n)$ of the model. This optimization can be conducted based on the expectation-maximization (EM) algorithm [4]. The EM algorithm maximizes the log-likelihood by iteratively applying an expectation step and maximization step until the calculation converges.

B. Experiments

To evaluate the validity of the inverse gamma-based variance distribution, EMG analysis experiments were conducted. Nine healthy young adults (males, right-handed, age: 23.4 ± 0.9 years) participated in the experiments. This study was conducted in accordance with the Declaration of Helsinki and was approved by the Hiroshima University Ethics Committee (Registration number: E-840).

During the experiments, the participants were seated with their right upper arm pointing downward, the right shoulder abducted at 45° , and the elbow flexed at 90° . The right forearm was pronated, and the fingers, except for the index finger, were fixed. The abduction force produced during isometric contraction of the first dorsal interosseous (FDI) muscle was measured using a force sensor (Leptrino, CF S018CA101U; 16-bit A/D; sampling frequency: 1,200 Hz) installed on the left side of the index finger. EMG signals were recorded from a wireless EMG sensor attached on the skin surface over the FDI muscle using a measurement system (Delsys, Trigno; 16-bit A/D; sampling frequency: 2,000 Hz; bandwidth: 20–450 Hz).

Before performing the EMG recording task, each participant performed at least two maximum voluntary contraction (MVC) trials. Additional MVC trials were performed if the difference between the peak force of the two MVCs deviated by over 5%. There was a rest period of at least 2 min between MVC trials. The largest peak force of all conducted MVC trials was taken as the MVC force. Then, the participants

were instructed to exert a constant isometric force at 5%, 10%, 15%, 30%, 50%, 70%, and 80% of the MVC force in a random order, and EMG signals were recorded simultaneously. The target force and exerted force were displayed on a monitor in front of the participants in real time. The participants were asked to maintain a constant force at the target level. Five seconds after starting to maintain muscle force, visual feedback of the exerted force was removed, and the participants were asked to maintain a constant effort for an additional 6 s. The latter 5 s of the recorded data without visual feedback was used for analysis. Three trials were conducted for each target force, and a minimum of 1 min of rest was taken between each measurement.

The variance distribution parameters α and β in the scale mixture model were estimated for each target force. To evaluate the validity of the inverse gamma-based EMG variance distribution, we calculated the empirical distribution of EMG variances. For this calculation, multiple variances needed to be calculated from a time-series EMG waveform from a single task. Therefore, we obtained EMG subsets by repeating simple random sampling with replacement from an EMG waveform from each task. Each subset contained W samples and random sampling was repeated 2,000 times. The variance value for each subset was then calculated, and the empirical distribution of the set of variances was obtained. To compare the model and empirical variance distributions, the Kullback–Leibler (KL) divergence [6] for both distributions was calculated as follows:

$$\text{KL}(p||q) = - \sum_{i=1}^m p(\sigma_i^2) \ln \frac{q(\sigma_i^2)}{p(\sigma_i^2)}, \quad (5)$$

where m is the sample size used for the calculation of the KL divergence and varied depending on data for each trial, and $q(\sigma_i^2)$ and $p(\sigma_i^2)$ are the model and empirical distributions, respectively. The KL divergence is a measure that represents the difference between two probability distributions; a KL divergence of zero indicates that the two distributions are identical. For the calculation of the KL divergence, the empirical distribution $p(\sigma_i^2)$ was calculated using kernel density estimation with a Gaussian kernel. In this experiment, the size of each subset for variance calculation was varied as follows: $W = 2, 3, 5, 10, 20, 50, 100, 200, 500$, and 1,000.

Then, the relationship between the model parameters and statistics of the empirical distribution was analyzed. Here, α and β can be assumed to represent the variation and scale of the EMG variance, respectively [4]. Therefore, regression analysis was performed between α and the coefficient of variation (CV) of the variance and between β and the mean of the variance. To investigate the relationship between the rate of change of each value, we conducted linear regression in log-log space (i.e., log-log linear regression). In this regression analysis, the value of W was determined based on the results of the KL divergence.

III. RESULTS

Fig. 2 shows examples of measured EMG signals and corresponding density histograms for the target forces of

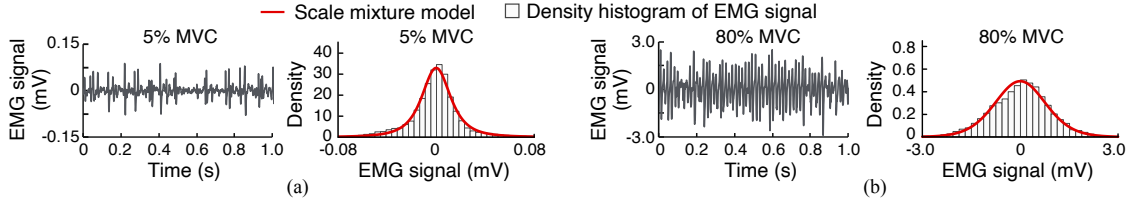


Fig. 2. Examples of measured EMG signals and corresponding density histograms. Fitted distributions based on the scale mixture model are also shown. (a) 5% MVC and (b) 80% MVC.

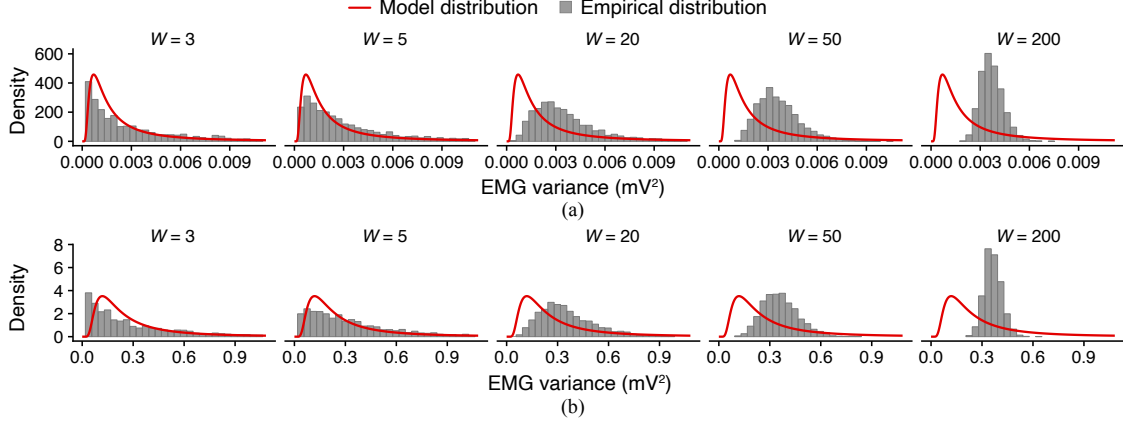


Fig. 3. Examples of the model and empirical distributions of the EMG variance. (a) 5% MVC and (b) 80% MVC. The sample size for the variance calculation in the empirical distribution was set to $W = 3$, $W = 5$, $W = 20$, $W = 50$, and $W = 200$.

5% MVC and 80% MVC. Fitted distributions with the scale mixture model are also shown as solid red lines. Fig. 3 shows examples of the model and empirical distributions of EMG variance for the target forces of 5% MVC and 80% MVC. In this figure, the sample size W for the variance calculation was set to 3, 5, 20, 50, and 200. The model distribution and empirical distribution are shown as solid red lines and gray histograms, respectively. Fig. 4 shows the relationship between the KL divergence and the sample size W for each target force. The vertical and horizontal axes were logarithmically scaled because of the wide-ranging values of both W and the KL divergence.

Regression analysis between the model parameters and the EMG variance statistics was performed in log-log scales (Fig. 5). Here, the sample size for the variance calculation was set to $W = 5$ based on the results of the KL divergence (Fig. 4). Fig. 5(a) shows the regression results between α and the CV of the EMG variance, and Fig. 5(b) shows the regression results between β and the mean of the EMG variance. The regression lines, regression equations, and coefficients of determination (R^2) are also shown in the figure. In both regression analyses, the model parameters and the EMG variance statistics were significantly correlated ($p < 0.0001$).

IV. DISCUSSION

The scale mixture model enables the description of changes in the shape of the EMG signal distribution according to muscle activity (as shown in Fig. 2). This is possible because the modeled variance distribution (i.e., inverse gamma distribution) also changes its shape and scale,

as shown in the examples of model distributions in Fig. 3. This figure shows that the empirical EMG variances tended to be distributed symmetrically when the sample size W was large. By contrast, as the sample size reduced, the distribution lost symmetry and had positive skewness; consequently, the empirical distribution approached the model distribution (i.e., inverse gamma distribution). This tendency can be observed for all target forces.

Such a change in the differences between the model distribution and the empirical distribution caused by the value of W was quantified using the KL divergence (Fig. 4). The KL divergence, which represents the difference between the distributions, decreased with a decrease in the sample size. This is consistent with the tendency observed in Fig. 3. Additionally, the KL divergence tended to have the minimum value at $W = 3$ or $W = 5$, on average. When the target force was small (5% and 10% MVC), the minimum value was observed at $W = 3$; by contrast, when the target force was higher than that, the minimum value was observed at $W = 5$. Thus, the sample size W at which the KL divergence was minimized depended on the target force level. This is thought to be because the shape of the variance distribution significantly varies between the large force level and the small force level [4].

The scale mixture model assumes that the EMG variance varies for each data point of EMG signals. Indeed, the difference between the model distribution and the empirical distribution became smaller as the sample size for calculating the empirical variance became smaller. These results potentially show the validity of the inverse gamma assumption for the EMG variance distribution. However, such inter-

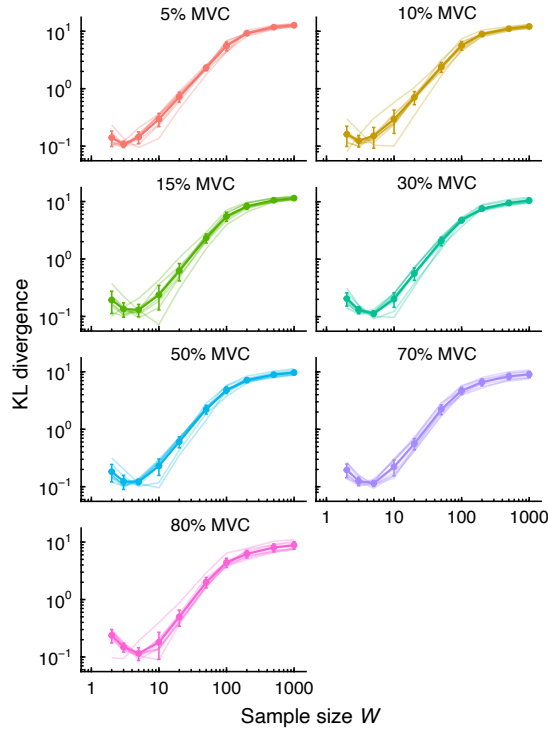


Fig. 4. Results of the Kullback–Leibler (KL) divergence between the model distribution and the empirical distribution with varied sample size W . Light colored lines represent the results of each participant. Error bars represent the standard deviations for all participants.

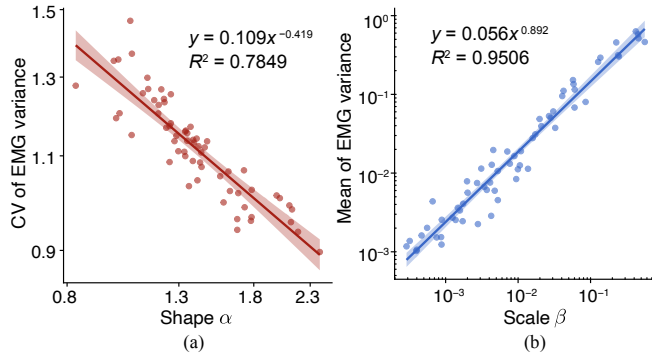


Fig. 5. Log-log regression analysis between the model parameters and the EMG variance statistics for all participants. (a) Shape α and CV of the EMG variance and (b) scale β and mean of the EMG variance. Solid lines represent fitted regression lines, and the regression equation and the coefficients of determination (R^2) are also shown. Shaded areas represent the 95% confidence interval around the regression lines.

distribution differences increased when the sample size was extremely small (for instance, $W = 2$). This is because EMG subsets for which the containing signal values became almost constant were likely to be generated if the sample size was too small, and the variance value calculated from such a subset may be smaller than the assumption of the model.

In Fig. 5, a linear correlation in log-log space was obtained for both the relationships between α and the CV of the variance and between β and the mean of the variance. In particular, a high coefficient of determination ($R^2 > 0.95$) was found in the regression between β and the mean of

the variance, which indicates that β strongly relates to the average scale of the EMG variance. By contrast, the authors previously assumed that the shape α , which characterizes the Gaussianity of the EMG, represents the relative variation of the EMG variance independent of scale [4]. The fact that a medium to high coefficient of determination ($R^2 = 0.7849$) was obtained between α and the CV of the variance—the relative variation of the variance—supports this hypothesis. Such a relative variation in the EMG variance reflects the proportion of outlier-like instantaneous spikes that are included in the EMG waveform, and it has been suggested that this is related to the recruitment mechanism of motor units [1], [4], [7].

From the above, the stochastic behavior of the EMG variance can be explained by the variance distribution parameters of the scale mixture model. The proposed model-based approach is considered to be more useful for investigating the stochastic property of the EMG variance than the empirical calculation-based approach. This is because the empirical approach requires setting the sample size and number of samplings to calculate the variance set, and the shape of the empirical distribution is affected by such settings. By contrast, the scale mixture model can evaluate the variance distribution without any direct calculation of variances.

V. CONCLUSION

In this paper, the validity of the variance distribution of EMG signals based on the scale mixture model was investigated. In the model, the EMG variance is considered as a random variable that follows an inverse gamma distribution.

In the experiments, the model distribution was compared with the empirical distribution, and the results potentially showed the validity of the inverse gamma assumption for the EMG variance distribution. Additionally, regression analysis showed that the stochastic behavior of the EMG variance can be explained by the parameters of the scale mixture model. In the future, the authors plan to examine the uncertainty of the EMG variance in more detail and demonstrate the relationship to motor noise and motor unit variability.

REFERENCES

- [1] H. S. Milner-Brown and R. B. Stein, “The relation between the surface electromyogram and muscular force,” *J. Physiol.*, vol. 246, no. 3, pp. 549–569, Apr. 1975.
- [2] E. P. Doheny *et al.*, “Effect of elbow joint angle on force–EMG relationships in human elbow flexor and extensor muscles,” *J. Electromyogr. Kinesiol.*, vol. 18, no. 5, pp. 760–770, Oct. 2008.
- [3] L. Mesin *et al.*, “Motor unit firing rates and synchronisation affect the fractal dimension of simulated surface electromyogram during isometric/isotonic contraction of vastus lateralis muscle,” *Med. Eng. Phys.*, vol. 38, no. 12, pp. 1530–1533, Dec. 2016.
- [4] A. Furui *et al.*, “A scale mixture-based stochastic model of surface EMG signals with variable variances,” *IEEE Trans. Biomed. Eng.*, vol. 66, no. 10, pp. 2780–2788, Nov. 2019.
- [5] H. Hayashi *et al.*, “A variance distribution model of surface EMG signals based on inverse gamma distribution,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 11, pp. 2672–2681, Nov. 2017.
- [6] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006, vol. 16.
- [7] D. F. Stegeman *et al.*, “EMG modeling and simulation,” in *Electromyography*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Jan. 2005, pp. 205–231.