

Exploiting Sparsity, Sparseness and Super-Gaussianity in Underdetermined Blind Identification of Temporomandibular Joint Sounds

Clive Cheong Took and Saeid Sanei

Centre of Digital Signal Processing, School of Engineering, Cardiff University, CF24 3AA, U.K.

Email: {cheongc, saneis}@cf.ac.uk

Abstract—In this paper, we study a 2×3 temporomandibular joint (TMJ) underdetermined blind source separation (UBSS). This particular UBSS has been subject to an empirical experiment performed previously on two *sparse* TMJ sources and a *non-sparse* source modelled as super-Gaussian noise. In this study, we found that FastICA algorithm tends to separate the two highly super-Gaussian sources when applied to the mixtures. When these two mixtures were filtered, FastICA focused on the non-sparse source (i.e. noise). Previously, we did not examine why such filtering approach would lead to estimation of the non-sparse source. To this end, the objective is to provide an extensive set of simulations to demonstrate why this filtering approach fully solve this particular underdetermined blind identification. We have employed the shape parameter α of the generalized Gaussian distribution (GGD) as a measure of sparseness and Gaussianity. This parameter was also utilized to illustrate the convergence of our filtering approach and the sub-Gaussian effect of the filter on the mixtures. Moreover, we have also considered the case where the noise source is modelled as sub-Gaussian and Gaussian as an extension of our previous work. Simulation studies show that our filtering approach is robust and performs well in this particular TMJ UBSS application.

Index Terms—sparsity, sparseness, Gaussianity, moving average filter, independent component analysis

I. INTRODUCTION

Blind source separation (BSS) is one of the most exciting current areas of research in statistical signal processing and unsupervised machine learning due to its potential applications in various areas such as financial time series analysis, biomedical signal processing, and digital communications [1], [2], [3]. The aim of BSS is to recover the sources from observations composed of mixtures, without a priori knowledge of the medium and the sources. Conventionally, BSS can be formulated as:

$$y_i(t) = \mathbf{w}_i^T \mathbf{V}(\mathbf{A}\mathbf{s}(t) + \mathbf{v}(t)) \quad (1)$$

This study investigates the role of sparsity in the empirical study entitled "A Filtering Approach to Underdetermined Blind Source Separation With Application to Temporomandibular Disorders," by Clive Cheong Took, Saeid Sanei, and Jonathon Chambers, which appeared in the Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, May, 2006, Toulouse, France.

where $y_i(t)$ denotes the i -th estimated source at discrete time instant t , \mathbf{w}_i^T the i -th row of the separating matrix \mathbf{W} , $(\cdot)^T$ the transpose operation, $\mathbf{s}(t)$ the $n \times 1$ source vector, $\mathbf{v}(t)$ the $n \times 1$ noise vector, and \mathbf{A} the unknown $n \times m$ mixing matrix. Furthermore, it is noted that we have access to only the $m \times 1$ mixture vector $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$ with the noise modelled as a third source as in [4].

Independent Component Analysis (ICA) is one approach to perform BSS for estimation of the sources $s_i(t)$, i.e. the independent components (ICs), assuming that the sources are statistically independent. FastICA [1] which figures amongst the most well-established algorithms in ICA, estimates the i th source by maximizing the non-Gaussianity of the estimated sources $\mathbf{y}(t)$. In this study, we will consider this algorithm. On the other hand if $n > m$, then the problem is coined as underdetermined BSS (UBSS). Firstly, it is required to *identify* \mathbf{A} , and thereafter source extraction is performed to solve fully UBSS. However, UBSS is quite challenging since 1) $n > m$ implies that there are less number of equations than variables, and therefore it is an ill-posed problem, and 2) no explicit a priori knowledge of \mathbf{A} and the sources is available. Therefore, many researchers have focused on the identification \mathbf{A} as in [5], [6], [7]. Likewise, one of the aims of this study is to investigate the role of sparsity in a 2×3 blind identification of \mathbf{A} . In particular, we consider the case where a pair of temporomandibular joint (TMJ) *sparse* sound sources prevail in the presence of a third *non-sparse* source modelled as noise as in [4]. In contrast to [4], we will also consider the possibility that noise can be modelled not only as a super-Gaussian source, but also as a sub-Gaussian or Gaussian source. The presence of a sub-Gaussian source in biomedical applications such as in [8] arises due to the proximity of the power line to the sensors, while it is known that noise is generally modelled as Gaussian [1], [9], [10].

In [4], we employed FastICA [1] to estimate the two columns of \mathbf{A} pertaining to the TMJ sources, while the third column of \mathbf{A} pertaining to the noise was estimated by a pre-filtering approach. However, this

approach was empirical, and we therefore propose to extend this work by analyzing various aspects of this particular TMJ BSS problem as follows. Hence, the objectives of this paper are to 1) demonstrate why linear filtering does not alter the structure of \mathbf{A} , 2) investigate the effect of considering different degrees of Gaussianity of the third noise source on the identification of \mathbf{A} , 3) demonstrate the relationship between sparsity and the degree of Gaussianity of the sources, and 4) illustrate how the pre-filtering approach leads to the identification of the third column of \mathbf{A} .

II. BACKGROUND AND RELATED WORK

A. The Temporomandibular Disorder BSS Problem

The temporomandibular joint is located between the mandible (lower jaw) and the temporal bone (skull). This pair of joints generate two types of sounds known as clicks and crepitus, when a person suffers from the temporomandibular joint disorder (TMD). The latter refers to all medical problems related to these two joints. The click is active for short and *distinct* periods, while the crepitus is a more noise-like signal, with approximately continuous active periods. These TMJ sources are illustrated in Fig. 1. The nature of the TMJ sound source depends on the type of TMD. For example, clicks arise as a result of the disc displacement, while a degenerative joint disease such as osteoarthritis leads to the generation of crepitus. A correct prognosis of TMD relies on the experience of the dental specialist and it has often caused controversy as in [11], and [12]. Likewise, diagnosis of TMD is particularly challenging when both types of TMJ sound sources prevail, in the presence of background noise as simulated in [4]. However, only a pair of mixtures can be recorded from a pair of stethoscopes placed inside the two auditory canals. This results in a 2×3 UBSS. However in this article, the main focus is on the blind identification of \mathbf{A} , as previously mentioned in introduction. Furthermore, we will vary the degree of Gaussianity of the noise source, instead of considering only the super-Gaussianity case as in [4]. Subsequently, we shall investigate how the degree of Gaussianity of this noise source can affect the identification of \mathbf{A} .

B. Sparsity, Sparseness and Super-Gaussianity

Sparsity (or disjointness) in this work refers to the situation where a relatively small number of source signals are active over any particular time interval. For the case of a single active source, sparsity [4] can be mathematically described as

$$\{s_i(t); i = 1, \dots, n\}$$

where $\forall t \exists k \in 1, \dots, n$ where $|s_k(t)| \gg |s_j(t)|$ (2)

and for $j \neq k$ $s_j(t) \approx 0$

where $s_i(t)$ is a given source signal and $s_j(t)$ is another source signal.

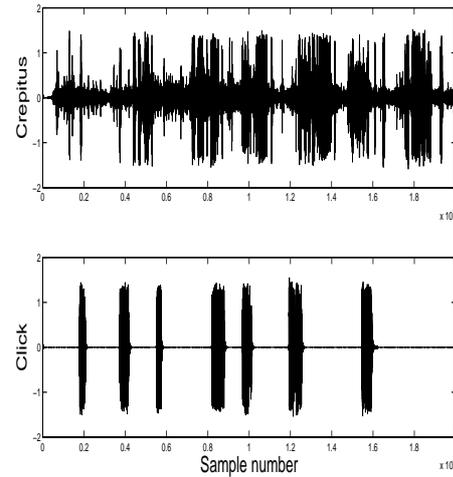


Fig. 1. The crepitus and the click signals. Note the sparseness (i.e. its tendency to be of zero magnitude) of the click compared to the crepitus.

On the other hand, the degree of sparseness of a source signal depends on the number of occurrences of its samples being zero or approximately zero. In [13], the authors refer to a sparse signal if the magnitude of most of its samples is zero or near zero, with only a few sample entries taking significant values. We note that this statement strongly correlates with the nature of an impulsive signal as in [9], because an impulsive signal consists of only a few high peaks of short duration. Cichocki classified a signal as impulsive if $0 < \alpha < 1$ (p. 245 [9]). On the other hand, He *et al.* categorize a signal to be sparse if its corresponding shape parameter α of the GGD is less than two [14]. As the shape parameter α of GGD seems to be a reasonable measure of sparseness, let us define such distribution as follows:

$$p(s_i(t), \sigma, \alpha) = \frac{\alpha\sqrt{\beta}}{2\sigma\Gamma(1/\alpha)} e^{-|\sqrt{\beta}s_i(t)/\sigma|^\alpha} \quad (3)$$

where $s_i(t)$ denotes the i -th source at discrete time t , $\sigma > 0$ is the scale parameter, $\Gamma(\cdot)$ is known as the gamma function, and $\alpha > 0$ is the shape parameter. As for $\sqrt{\beta} = \sqrt{\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}}$, it is merely a scaling factor which enables $\text{var}(s_i(t)) = \sigma^2$ where $\text{var}(\cdot)$ denotes variance. A signal is said to be Gaussian distributed if $\alpha = 2$. However when the shape parameter α of a signal is less than two, it has a super-Gaussian or leptokurtic distribution, while with $\alpha > 2$, it has a sub-Gaussian/platykurtic distribution [1]. Hence, He *et al.* suggested the equivalence between a super-Gaussian signal and a sparse signal. However, the strong similarity between the impulsiveness definition and that of sparseness given by Pearlmutter [13] suggests that the latter was more restrictive in his definition of sparseness, i.e. $0 < \alpha < 1$. In effect, for the sparsity condition (2) to be fulfilled, the sparseness definition of Pearlmutter seems more appropriate. This is because as α tends to zero, the probability of the signal to be of zero magnitude increases as was shown in [14]. Therefore, we will call a signal sparse if $0 < \alpha < 1$.

$(\mathbf{A}^T \mathbf{A})$ is non-singular and therefore invertible. Hence, $\mathbf{s}(t) = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{x}(t)$ and $\mathbf{q}(t) = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{x}(t)$ are equal, which concludes the proof.

Theorems 10.3.5 and **10.3.8** which imply the non-Gaussianity assumption **A1** together with the implicit knowledge of the number of sources, and the linear independence of the columns of \mathbf{A} guarantees the uniqueness of the model $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$. However, it should be stressed that the model $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$ is still unique if there is at most one Gaussian source. This statement follows from the corollary of **theorem 10.3.6** [15] and from **theorem 5** of [16].

On the other hand, **A3** implicitly implies **A4** which will be explained as follows. The kurtosis $\text{kurt}(\mathbf{s}_i)$ which is a measure of the ‘peakedness’ of the probability distribution of \mathbf{s}_i can be defined as [17]:

$$\text{kurt}(\mathbf{s}_i) = \frac{E\{\mathbf{s}_i^4\}}{E\{\mathbf{s}_i^2\}^2} - 3 \quad (9)$$

where $E\{\cdot\}$ stands for the expected value. If $\text{kurt}(\mathbf{s}_i) = 0$, then \mathbf{s}_i is referred to as Gaussian, while $\text{kurt}(\mathbf{s}_i) > 0$ implies that \mathbf{s}_i has a super-Gaussian distribution. Otherwise, \mathbf{s}_i is known as sub-Gaussian. As the r th moment of a GGD can be expressed as follows [18]:

$$E\{|\mathbf{s}_i|^r\} = \frac{\Gamma(\frac{r+1}{\alpha})}{\Gamma(\frac{1}{\alpha})} \beta^{r/2} \quad (10)$$

We can then formulate (9) in terms of α using (10) as:

$$\text{kurt}(\mathbf{s}_i) = \frac{\Gamma(\frac{5}{\alpha})\Gamma(\frac{1}{\alpha})}{\Gamma^2(\frac{3}{\alpha})} - 3 \quad (11)$$

Thus we can plot $\text{kurt}(\mathbf{s}_i)$ as a function of α as shown

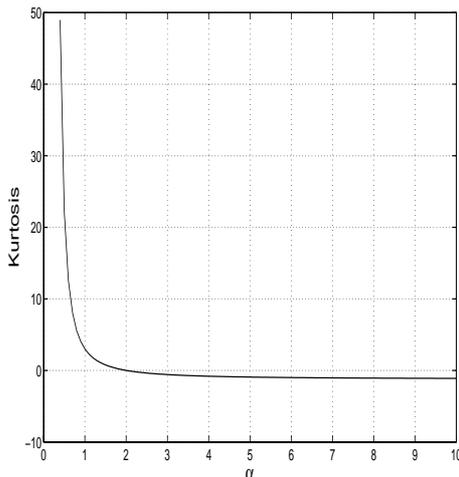


Fig. 2. Kurtosis as a function of α . Note that when $\alpha > 2$, kurtosis < 0 for the sub-Gaussian case and the rate of change of kurtosis is much lower than that of the super-Gaussian case ($\alpha < 2$)

in Fig. 2. Notice that the kurtosis has a large value when $0 < \alpha < 1$. This in turn implies the highly super-Gaussianity of \mathbf{s}_i for $0 < \alpha < 1$. Recall that we

classify a source signal as sparse if $0 < \alpha < 1$. Thus, as **A3** states that there are two highly super-Gaussian sources, it implies their sparseness (**A4**) as well. Hence, we stress that the equivalence between **A3** and **A4** is not straightforward without the formulation of equation (11). As FastICA maximizes the non-Gaussianity of the estimated ICs, this algorithm will focus on the two-highly super-Gaussian sources (**A3**) as will be shown in section III-B.

In the context of blind identification of \mathbf{A} , the resumé of the underdetermined UBSS FastICA is given below:

- FastICA is applied to \mathbf{X} to compute two columns of \mathbf{A} pertaining to the two highly super-Gaussian sources. The key point here is that FastICA will focus on the highly super-Gaussian sources due to the maximization of negentropy in (4).
- \mathbf{X} is pre-filtered as shown in (5) to yield \mathbf{X}' . The inputs to the FastICA are the resulting \mathbf{X}' to estimate the third column pertaining to the noise. The column of \mathbf{A}' corresponding to the IC with the lowest kurtosis is selected as the third column of \mathbf{A}' due to assumption **A3**.

It is understood that the two estimated columns of \mathbf{A}' can be computed by the inversion of the separating matrix $\mathbf{W}' = [\mathbf{w}_1 \mathbf{w}_2]^T$ of FastICA to be used in each of the above two steps.

III. CASE STUDY

A. Simulation

In this section, we investigate how the filtering UBSS FastICA approach [4] performs subject to different conditions via extensive simulation studies. The conditions we have considered were: the Signal-to-Noise Ratio (SNR), the degree of Gaussianity of the noise, and the filter length. The three distributions utilized in this study was Laplace, normal, and uniform to convey respectively the super-Gaussianity, Gaussianity and sub-Gaussianity nature of the non-sparse source. Furthermore, we have monitored the evolution α to assess the convergence of the UBSS FastICA. The performance measure (PM) that provides an indication of the difference between \mathbf{A} and its estimated $\hat{\mathbf{A}}$ is employed here [19]. However, PM requires both \mathbf{A} and $\hat{\mathbf{A}}$ to have unit norm columns. This index falls within $0 \leq PM \leq 1$. PM equals to 0 if $\hat{\mathbf{A}} = \mathbf{A}\mathbf{P}$ where \mathbf{P} is a permutation matrix. Therefore the lower the PM , the better is the performance of the UBSS algorithm.

$$PM(\mathbf{A}, \hat{\mathbf{A}}) = 1 - \left(\frac{1}{2n} \sum_{i=1}^n \sup_j |\mathbf{A}^T \hat{\mathbf{A}}|_{ij} + \frac{1}{2n} \sum_{j=1}^n \sup_i |\mathbf{A}^T \hat{\mathbf{A}}|_{ij} \right) \quad (12)$$

For each simulation where one condition was varied, 20 independent Monte Carlo trials were run and averaged to provide the graphs in Fig. 3-8. However, before we proceed to assess the performance of UBSS FastICA, we investigate the effect of the filter length on the mixtures $\mathbf{x}(t)$ in terms of their Gaussianity via α . This is illustrated

in Fig. 3.

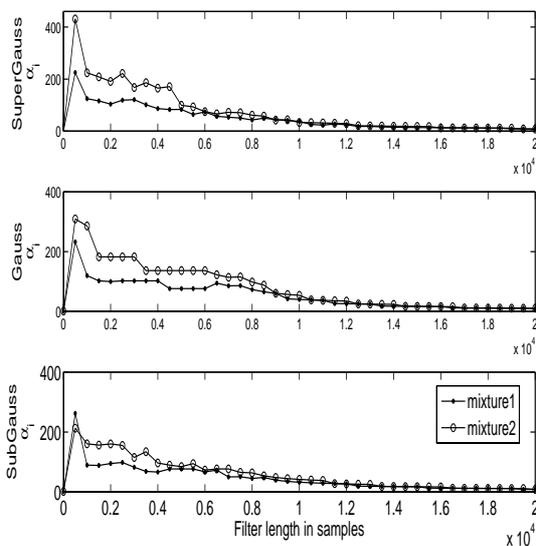


Fig. 3. Effect of filter length on the degree of Gaussianity of the mixtures when super-Gaussian, Gaussian, sub-Gaussian noise (from top to bottom) at 0 dB. It is noteworthy to say at the maximum filter length $M = 20\ 000$ samples, $\alpha > 5$. In other words, the mixtures are still sub-Gaussian. However, prior to pre-filtering of the mixtures, $\alpha < 1$. This explains why without filtering, FastICA focuses on the TMJ sources, while pre-filtering leads to the non-sparseness of the mixtures and consequently estimate the non-sparse noise instead of the sparse TMJ sources.

In Figs. 4 & 5 we assess the performance of UBSS FastICA, as SNR and filter length were varied. Note that for each simulation, we have considered noise as the non-sparse source whose distribution ranged from super-Gaussian to sub-Gaussian.

Thereafter, we simulate the 2×3 TMJ UBSS at SNR=0 dB when the degree of Gaussianity of the noise was altered from Figs. 6, 7, and 8. From these plots, we show the convergence in terms of α of the estimated ICs at each iteration. The true values α of the ICs are also included.

B. Discussion

In the first place, we note the sub-Gaussianity effect on the mixtures $\mathbf{x}(t)$ by moving average filtering in Fig. 3. However prior to filtering, α of the mixtures was less than unity and therefore the signals were highly super-Gaussian and sparse. In [4], we intuitively stated that the filtering process suppressed the two highly super-Gaussian sources in the mixtures and therefore the *non-sparse* noise was more prominent in $\mathbf{x}(t)$. In section II-B, it was deduced that when $\alpha \geq 1$, a signal is non-sparse. On this basis, we re-affirm that the filtering suppress the two highly super-Gaussian *sparse* sources. This is because the moving average has altered the nature of the mixtures from highly super-Gaussian

to sub-Gaussian or equivalently from sparse to non-sparse. On the other hand, it is clear that as the filter length M increases, both mixtures tend to have Gaussian distribution due to the central limit theorem. Nonetheless, it is worth noting that at the maximum filter length of $M = 20\ 000$ samples, $\alpha > 5$ indicates that the mixtures are still sub-Gaussian.

With regard to the PM against SNR shown in Fig. 4, notice the maximum value of PM is of order 10^{-3} . This demonstrates the good performance of UBSS FastICA in all the three scenarios. Observe the much better performance of the sub-Gaussian case. This was expected as further to our discussion in the previous paragraph, where we addressed the sub-Gaussian/ non-sparse effect of the filter on the mixtures. Additionally, all three curves have a minimum at SNR=10 dB. However, this is not obvious for the sub-Gaussian case from Fig. 4 due to its much lower performance measure (of magnitude of order 10^{-5}). The minimum at 10 dB arises due to the following: the higher the SNR, the better the estimation of the two columns of \mathbf{A} pertaining to the two TMJ sources and *vice-versa* for the noise. These explain the best performance mid-way between 0 dB and 20 dB.

Next, we discuss the performance of UBSS FastICA in terms of the filter length in Fig. 5. However before we proceed further, let us go back to Fig. 3. This graph might be misleading in terms of the *high* magnitude of α in the sense of the degree of sub-Gaussianity. Fig. 2 demonstrates that for the sub-Gaussian case (i.e. $\alpha > 2$), the kurtosis does not change exponentially as when $\alpha < 2$. For example, the difference between the kurtoses corresponding to $\alpha = 3$ and $\alpha = 400$ is not

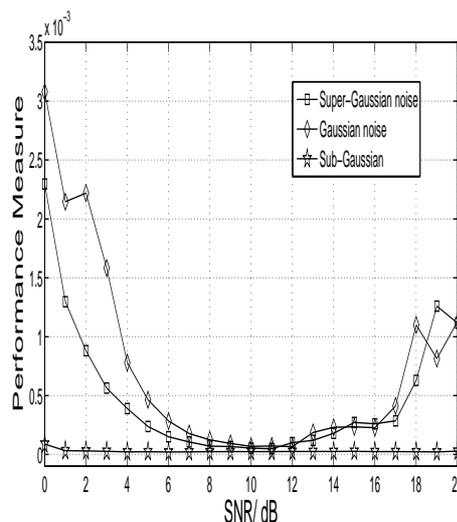


Fig. 4. Performance Measure versus Signal-to-Noise-Ratio (SNR) in dB when super-Gaussian, Gaussian and sub-Gaussian noises were considered. Note the much better performance measure of the sub-Gaussian noise case. This is because pre-filtering leads to the non-sparseness/sub-Gaussianity of the mixtures as seen in Fig. 3. Therefore their distributions are much closer to the sub-Gaussian noise.

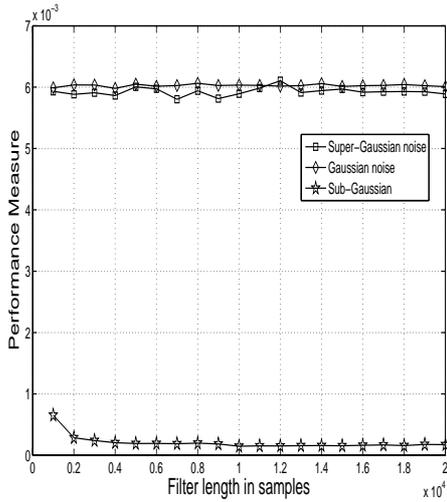


Fig. 5. Performance Measure as a function of filter length at 0 dB when super-Gaussian, Gaussian and sub-Gaussian noises were considered. Note the much better performance measure of the sub-Gaussian noise case. In fact, the moving average pre-filtering leads to the non-sparseness/sub-Gaussianity of the mixtures as seen in Fig. 3 . Therefore the distributions of the mixtures are much closer to the sub-Gaussian noise.

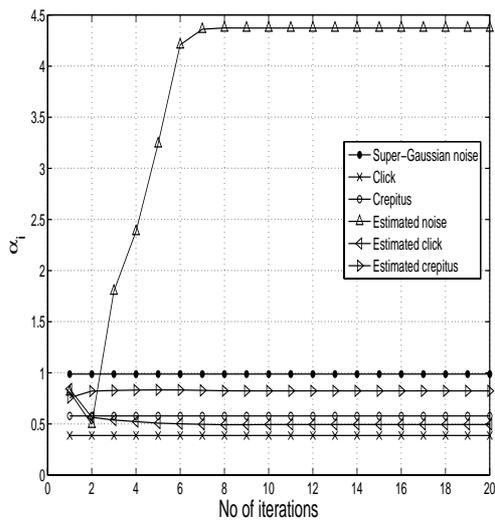


Fig. 6. Convergence graph: Evolution of α at 0 dB when super-Gaussian noise was considered.

significant (i.e. difference ≈ 0.6). Hence, the degree of sub-Gaussianity from $\alpha > 2$ to $\alpha = 400$ does not change significantly. Hence, the filter length does not alter significantly the degree of sub-Gaussianity of the mixtures. From Fig. 5, we note that the filter length does not have any significant effect on the performance measure (all three curves are approximately unvaried). Further to our previous discussion, such trend was expected. Moreover, the superior performance of the sub-Gaussian case is again highlighted.

Last but not least, we examine the convergence of UBSS FastICA in terms of α when super-Gaussian,

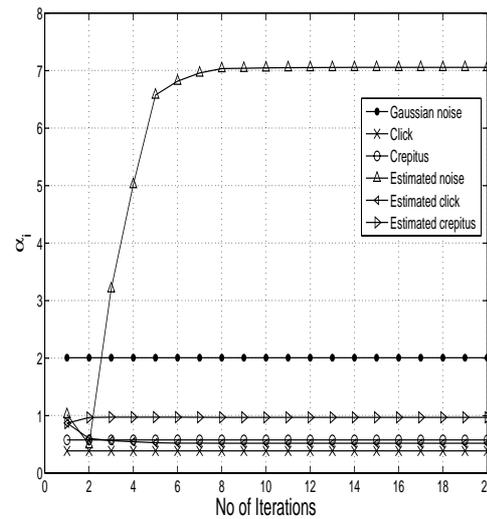


Fig. 7. Convergence graph: Evolution of α at 0 dB when Gaussian noise was considered.

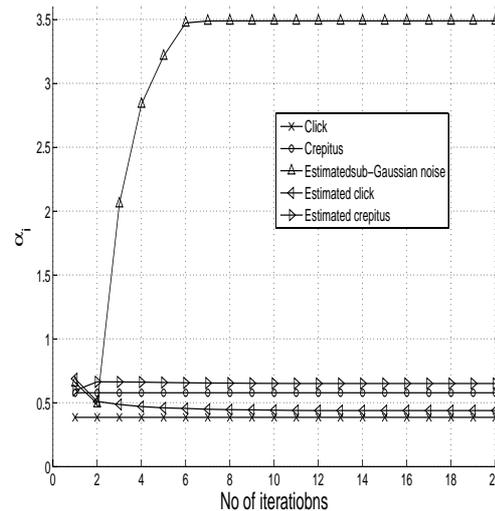


Fig. 8. Convergence graph: Evolution of α at 0 dB when sub-Gaussian noise was considered. Note the closeness of the α_i of the estimated TMJ sources with those of the original TMJ, compared to the super-Gaussian and Gaussian noise cases in Fig. 6 & 7 respectively.

Gaussian and sub-Gaussian noises considered in Figs. 6, 7, & 8 respectively. From these plots, the most striking curve (topmost) belongs to the estimated noise. From Figs. 6 & 7, the estimated noise converges to sub-Gaussianity as $\alpha > 2$. This contradicts the super-Gaussianity and Gaussianity nature of the noise considered. Nevertheless if we look at equation (5), we can deduce that we are estimating the filtered ICs S' , instead of the original S . As discussed previously, the moving average process in this particular TMJ UBSS accounts for this sub-Gaussian/ non-sparse effect. On the other hand, if we inspect the convergence of the estimated TMJ ICs (i.e. click and crepitus), we can see that in all cases they converge to $\alpha \leq 1$ (i.e. highly

super-Gaussianity). The reason why they do not converge to the true values is probably due to the low SNR of 0 dB. Again, we emphasize the much better performance of sub-Gaussian case in Fig. 8. This is illustrated by the closeness of the estimated TMJ ICs to the true ones in Fig. 8 compared to those in Figs. 6 & 7.

IV. CONCLUSION

In this paper, we have investigated the role of sparsity in the empirical study in [4] via the shape parameter α of the generalized Gaussian distribution. We have seen a close relationship between this parameter, the sparseness of a signal and consequently the sparsity situation in this particular TMJ UBSS. We have implicitly explained the subtle difference between sparsity and sparseness. We have gathered that a highly super-Gaussian signal, i.e. $\alpha < 1$ is likely to be sparse in section II-B. Furthermore, we have demonstrated why the ICA model still stands after pre-filtering in equation (5) in terms of the statistical independence of the new IC S' and the unaltered structure of A .

In the simulation studies, we deemed that the moving average filter has a sub-Gaussian/non-sparse effect on these particular TMJ mixtures $x(t)$. Consequently, the sub-Gaussian noise TMJ UBSS outperformed the other scenarios, i.e. when super-Gaussian and Gaussian noises were considered. However, the performances in all three scenarios were good due to their performance measure being of order 10^{-3} or less. Equation (5) demonstrates why the estimated noise does not converge to the original ones in Figs. 6 & 7. In [4], we intuitively stated that the filter suppressed the two *sparse* TMJ sources. In the sequel, the *non-sparse* noise pre-dominates in the filtered mixtures. This is obvious from the *sub-Gaussian/non-sparse* effect of the filter in Fig. 3.

REFERENCES

- [1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. John Wiley & Sons, INC, 2001.
- [2] S. Sanei and J. Chambers, *EEG Signal Processing*. John Wiley, 2007.
- [3] T.-W. Lee, *Independent Component Analysis : Theory and Applications*. Boston : Kluwer Academic Publishers, 1998.
- [4] C. Cheong, S. Sanei, and J. Chambers, "A Filtering Approach to Underdetermined Blind Source Separation with Application to Temporomandibular Disorders," *ICASSP 06 Proceedings, IEEE International Conference*, pp. III-1124 – III-1127, May 2006.
- [5] P. Comon and M. Rajih, "Blind Identification of Under-Determined Mixtures Based on The Characteristic Function," *Elsevier Signal Processing*, vol. 86, pp. 2271–2281, Oct 2006.
- [6] J. F. Cardoso, "Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors," *Proceedings of the ICASSP, Toronto*, pp. 3109–3112, 1991.
- [7] A. Taleb, "An algorithm for the blind identification of N independent signals with 2 sensors," *IEEE International Symposium on Signal Processing and its Applications*, pp. 5–8, August 2001.
- [8] M. G. Jafari and J. A. Chambers, "Fetal electrocardiogram extraction by sequential source separation in the wavelet domain," *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 390 – 400, Mar 2005.
- [9] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing- Learning algorithms and Applications*. John Wiley and Sons Ltd, 2005.
- [10] S. M. Kay, *Fundamentals of Statistical Signal Processing- Estimation Theory*. Prentice Hall Signal Processing Series, 1993.
- [11] R. J. M. Gray, S. J. Davies, and A. A. Quayle, *Temporomandibular Disorders: A Clinical Approach, 1st Ed.* British Dental Association, 1995.
- [12] C. S. Greene, N. D. McNeill, C. Clark, and G. T. Truelove, "Temporomandibular disorders and science: a response to the critics," *Journal of Prosthetic Dentistry*, vol. 80, p. 214, 1998.
- [13] B. A. Pearlmutter and V. K. Potluru, "Sparse Separation: Principles and Tricks," *Proceedings of SPIE, Independent Component Analyses, Wavelets, and Neural Networks*, vol. 5102, pp. 1–4, April 2003.
- [14] Z. He, S. Xie, and Y. Fu, "Sparseness Measure of Signal," *IEEE International Conference on Neural Networks and Brain, ICNN&B '05*, vol. 3, pp. 1931 – 1936, Oct 2005.
- [15] A. M. Kagan, I. Linnik, and C. R. Rao, *Characterization Problems in Mathematical Statistics*. John Wiley & Sons Inc, 1973.
- [16] X.-R. Cao and R.-W. Liu, "General Approach to Blind Source Separation," *IEEE Transactions on Signal Processing*, vol. 44, pp. 562–571, Mar 1996.
- [17] M. Kendall, A. Stuart, and K. Ord, *Kendall's Advanced Theory of Statistics: Distribution Theory*. Hodder Arnold; 6th Revised Edition, 1994.
- [18] H. Mathis and S. C. Douglas, "On the existence of universal nonlinearities for blind source separation," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1007–1016, May 2004.
- [19] F. C. Meinecke, S. Harmeling, and K. R. Muller, "Robust ICA for Super-Gaussian Sources," *Independent Component Analysis and Blind Source Separation: Fifth International Conference*, vol. 5, pp. 217–224, 2004.

Clive Cheong Took is currently a Ph.D. candidate within the Centre of Digital Signal Processing, at Cardiff University, UK. He received his Bachelor degree in Telecommunication Engineering from Kings College London University, in 2004. His research interests include source separation, adaptive and blind signal processing.

Saeid Sanei received his PhD from Imperial College of Science, Technology, and Medicine, London, in Biomedical Signal and Image Processing in 1991. He has been a member of academic staff in Iran, Singapore, and United Kingdom. His major interest is in biomedical signal and image processing, adaptive and nonlinear signal processing, and pattern recognition and classification. He has had a major contribution to Electroencephalogram (EEG) analysis and brain computer interfacing (BCI). Currently, he is the member of staff in the Centre of Digital Signal Processing, Cardiff University, United Kingdom, and is a Senior Member of IEEE.