

Analysis of spike train data: A Multivariate Mixed Effects Model for Phase and Amplitude

Pantelis Z. Hadjipantelis^{1,2}, John A.D. Aston¹

1. Centre for Research in Statistical Methodology, University of Warwick, UK

2. Centre Complexity Science, University of Warwick, UK

e-mail: p.z.hadjipantelis@warwick.ac.uk; j.a.d.aston@warwick.ac.uk

Hans-Georg Müller

Department of Statistics, University of California, Davis, USA

e-mail: hgmueLLer@ucdavis.edu

John Moriarty

School of Mathematics, University of Manchester, UK

e-mail: j.moriarty@manchester.ac.uk

Abstract: This discussion of the spike train data applies the approach of Hadjipantelis *et al.* developed for Linguistic data to jointly model both the phase and amplitude functions via a mixed effects model. The approach shows that care needs to be taken when assessing amplitude and phase functions, particularly if separate analysis is undertaken, as there can be high levels of correlation between the warping and amplitude component functions.

Keywords and phrases: Multivariate mixed model, Functional mixed effects regression, Pairwise warping.

1. Introduction

When a monkey conducts a task, brain activity can be measured using single electrodes connected to the neurons of its brain. Each electrode records changes in the action potential of neurons and these sequences of fluctuating action potentials are what we refer to as spike train data. The data set and introduction to spike trains is provided in Wu, Hatsopoulos & Srivastava [1]; background is given in [2]. An important characteristic of a given spike train waveform is that the amplitude of the fluctuations is relatively uninformative while the temporal placement of the spikes, their phase [3], carries the most important information. Our data are 240 trials of a hand movement task; the movement pattern outlines the perimeter of a square. The subject had to move a cursor from one angle to the next angle in a counter-clockwise fashion and the task was completed when the cursor returned to each original starting angle. The experiment had a balanced design with 60 runs per path and all waveforms were normalized to the same lengths [2].

Spike train data present an obvious test-bed for methodologies that can concurrently account for length, amplitude and phase variations. As shown in other studies [3] a pre-registration step significantly enhances the explanatory power of modelling spike trains, as it allows more meaningful interpretation of covariance patterns in conjunction with the

registration functions. We are interested in determining whether more relevant information can be extracted from spike train data by jointly considering amplitude and phase, rather than considering these components separately, providing an illustration of the statistical methodology of Hadjipantelis et al [4].

While a number of studies have focused on the analysis of electrophysiological responses as functional data and even more specifically as functional data embedded in a random-effect structure [5, 6, 7, 8], these studies employ functional statistical methodology exclusively for the amplitude domain. They assume that inference can be conducted directly on the spike waveforms (or a smoothed version of them). We recognise this as an obvious limitation and present a study that accounts for variations in amplitude and phase domains concurrently.

This discussion aims to address three main questions:

- What are the principal modes of variation in the sample when one accounts for phase variability?
- Is there significant correlation between phase and amplitude variations and what consequences does any correlation have?
- Can we produce exemplar spike train waveforms that correspond to our initial dynamic assumptions?

2. Statistical Methodology

Our proposed statistical framework is based on methodology outlined in [4]. As in Tang & Müller [9] we view the i th waveform f_i , measured over the interval $t \in [0, 1]$, as :

$$f_i(t) = \tilde{f}_i(h_i^{-1}(u)). \quad (1)$$

A waveform $f_i(t)$ is viewed as the realization of an amplitude variation function \tilde{f} evaluated over u , with the mapping $h(\cdot)$ transforming the real time t onto the universal/sample-wide time-scale u . We also assume that there is a random effect Z_i associated with each curve, accounting for the path taken in the particular experimental condition.

\tilde{f}_i dictates the size of a given feature and h_i^{-1} dictates the location of that feature. We consider a representation of both the amplitude and phase functions as basis expansions:

$$\tilde{f}_i(u) = \mu^{\tilde{f}}(t) + \sum A_i^{\tilde{f}} \zeta_i^{\tilde{f}}(u), \quad \tilde{h}_i(u) = \sum A_i^{\tilde{h}} \zeta_i^{\tilde{h}}(u), \quad (2)$$

where $\mu^{\tilde{f}}(t)$ is the mean of the amplitude functions and where

$$h_i(t) = \frac{\int_0^t e^{\tilde{h}_i(u)} du}{\int_0^1 e^{\tilde{h}_i(u)} du} \quad (3)$$

Clearly different choices of bases $\zeta^{\tilde{f}}, \zeta^{\tilde{h}}$ will give rise to different coefficients A , which can be used for subsequent analysis such as functional regression as illustrated below; a number of different parametric basis functions can be used as bases [10]. Here, we advocate the use of a principal component basis both in the case of \tilde{f}_i and \tilde{h}_i , using $\tilde{h}_i(u)$ rather than $h_i(u)$ directly so as to yield a valid space to perform FPCA, as principal components provide a parsimonious representation in terms of variance explained [11].

We note that model (Eq. 2) is not identifiable in its most general form, and a solution adopted in [9] is to assume that the variation in the amplitude functions \tilde{f}_i will asymptotically vanish, whence the model can be shown to be identifiable under relatively weak additional regularity conditions. Practically, this can be justified by normalizing the amplitude functions and assuming that this removes most the variation in these functions. This approach is particularly suited for situations where the variation in the registration functions is more important than that of the amplitude functions, which is well justified for spike train data as already mentioned above, or where one uses pre-specified bases for both amplitude and registration functions. If the registration functions are always determined first, the estimates will also be unique, as will the amplitude estimates, but in case of non-identifiability, this could cause difficulties in interpretation.

As mentioned above, we assume that each waveform component is subjected to path-specific variational pattern; we incorporate this information in the covariate Z_i , which takes one of four categories depending on the path. As such the general form of our model for a given sample curve f_i with scalar covariates Z_i is:

$$E\{\tilde{f}_i(u)|Z_i\} = \mu^{\tilde{f}}(t) + \sum_k E(A_{i,k}^{\tilde{f}}|Z_i)\phi_k(u), \quad (4)$$

and

$$E\{\tilde{h}_i(u)|Z_i\} = \mu^{\tilde{h}}(u) + \sum_m E(A_{i,m}^{\tilde{h}}|Z_i)\psi_m(u). \quad (5)$$

$\mu^{\tilde{f}}(t)$ is estimated by the sample mean on the amplitude variation functions \tilde{f} and $\mu^{\tilde{h}}(u)$ is estimated by the sample mean of the phase variation functions \tilde{h} , but which will be close to zero as the warpings are designed such that the expectation is that no warping occurs, corresponding to a zero mean.

Following the registration of the original waveform data f we produce the registered waveforms \tilde{f} and their corresponding time-registration functions \tilde{h} . Then fixing the set of basis functions ϕ and ψ as the functional principal components of the amplitude and the phase covariance functions respectively, we estimate the scores $A_i^{\tilde{f}}$ and $A_i^{\tilde{h}}$. Those scores will act as surrogate data for curve f_i and can be jointly modelled via a pure multivariate random effects model as formulated below:

$$E([A_{i,k}^{\tilde{f}}, A_{i,m}^{\tilde{h}}]|Z_i) = Z_i\Gamma \quad \Gamma \sim N(0, \Sigma_\Gamma), \quad (6)$$

where $[A_{i,k}^{\tilde{f}}, A_{i,m}^{\tilde{h}}]$ is the $N \times (k+m)$ matrix of FPCA scores, where k is the number of amplitude and m the number of phase components retained. The $(k+m) \times (k+m)$ matrix Σ_Γ is the covariance matrix of the associated random effects. The role of Z_i , the design matrix for path-specific effects, is to allow different intercepts among the different path groups. A much more detailed and comprehensive discussion of this modelling framework and computation is presented in [4]. We note that multivariate covariates Z_i can also be incorporated by using single index or additive modelling approaches for the conditional expectations in Eq. 4 and 5.

The model above depends on the joint phase and amplitude variation functions available. To estimate phase variation/warping functions we take the approach of Tang & Müller [9], as implemented in the routine WFPCA in PACE. We first define a pairwise warping function $g_{i',i}$ between two sample curves, where the pairwise warping function $g_{i',i}$

acts as a 1-to-1 mapping from the i -th curve time-scale to that of the i' -th. The inverse of the sample-wide temporal mapping for f_i, h_i^{-1} , is then defined as the MLE of the $g_{i',i}$ sample, i' being an index over a sufficiently large sub-sample of curves from our original sample. Different methodologies based on the square-root velocity function metric [12] or area under the curve normalization [13] could be utilized to give a different, but interchangeable in subsequent analysis, set of warping functions.

3. Data Analysis and Results

Our first insights from the sample come directly from the warped data themselves (Fig. 1); it is evident that we have localized variation associated with the identity of the path examined. The movement that produces this “increase” is the one from point 4 to point 1, suggesting that the vertical upward movement pattern excites the neurons of the subject monkey primary motor cortex. It can be argued that a “partial” neuron excitement also occurs on the movement from 1 to 2.

A qualitative examination of the FPCA generated modes of principal variations suggests, that as expected, the amplitude FPCs have no single dominant variational tendency (Appendix: Table 3, Column: Amplitude). We expected

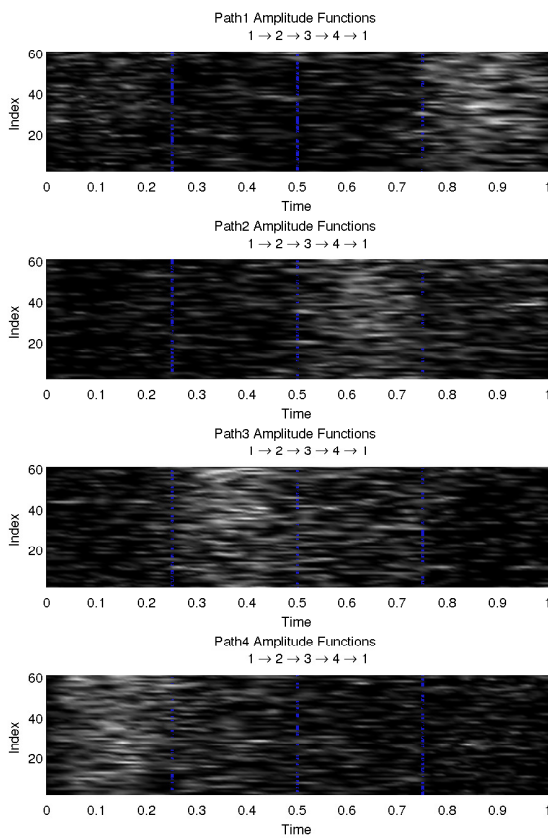


FIG 1. The four different warped waveform groups \tilde{F} . Blue lines are added to signify the change from move $x_i \rightarrow x_{i+1}$ to move $x_{i+1} \rightarrow x_{i+2}$.

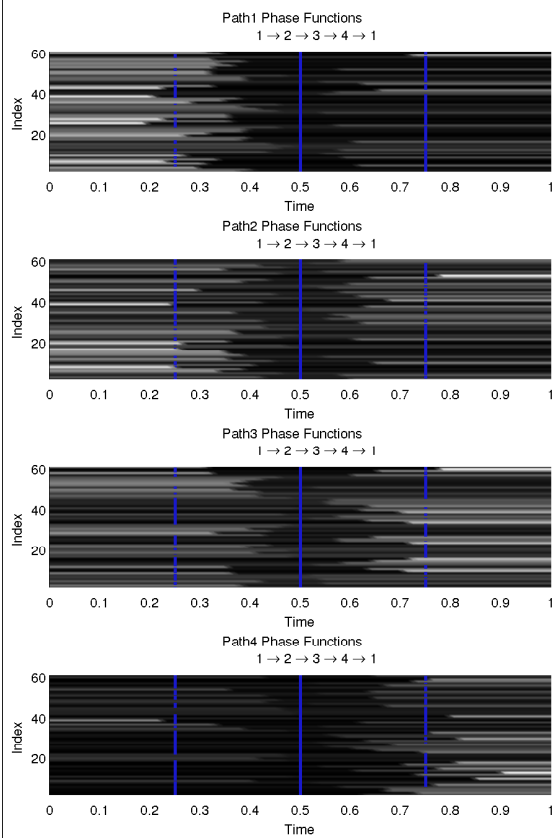


FIG 2. The four different groups of phase functions \tilde{H} . Blue lines are added to signify the change from move $x_i \rightarrow x_{i+1}$ to path $x_{i+1} \rightarrow x_{i+2}$.

this as we have four distinct motion patterns that invoke four distinct activity patterns; this being also supported by the plot of means, which hints at the presence of four peaks (Fig. 3, upper left plot). Additionally the shapes of the amplitude FPCs show rather abrupt changes instead of smooth varying ones and further suggest partially unrelated patterns of variation between the samples. In particular the first amplitude FPC shows a strong negative correlation between paths 4 & 3 and paths 2 & 1. Moreover the same qualitative behaviour with roughly equal intensity can be seen in second amplitude FPC between paths 4 & 1 and paths 3 & 2.

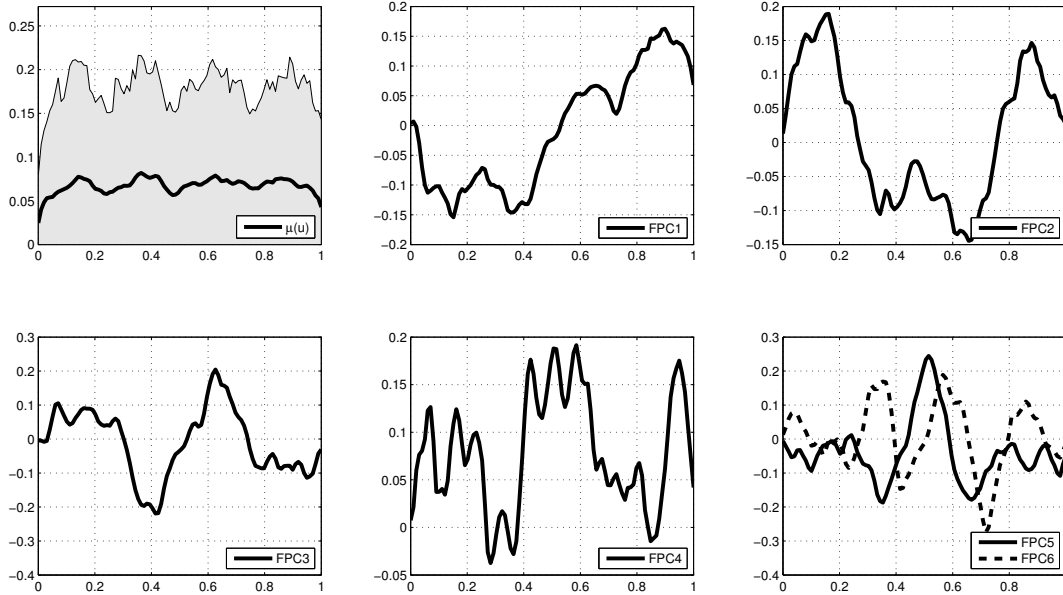


FIG 3. Mean function (.1 & .9) quantiles shown in grey area) and first, second, third, fourth, fifth, and sixth functional principal components. Together these account for 75.20% of the sample variance, but only the first four are considered informative (63.08 % of samples variation) and as such the fifth and sixth were not used in the subsequent analysis.

It is quite interesting that these two components appear to be able to fully qualitative describe all major path behaviours. Path1 can be approximately constructed as being a positive linear combination of FPC1 and FPC2, Path2 as being positively influenced by FPC1 and negatively influence by FPC2, Path3 as being a negative linear combination of FPC1 and FPC2, and Path4 as being negatively influenced by FPC1 and positively influenced by FPC2. Encouragingly, these intuitive assumptions are supported by the examination of the realizations of the random effect (Table 1, Columns: $\tilde{f}FPC1$, $\tilde{f}FPC2$) where the signs of the vectors $\gamma_{\tilde{f}FPC1}$ and $\gamma_{\tilde{f}FPC2}$ convey exactly the same intuition. On the contrary amplitude FPC3 seems to embody a localized counter-correlation pattern between Paths 3 and 2; amplitude FPC4 showing a similar but rather noisier pattern again focusing primarily on the distinct behaviour of Path2 compared to Path3. We feel that while “only” 63% of the total amplitude variance is retained using the first four FPCs, given the qualitative characteristics of those FPCs, they are enough to adequately represent all the main transitions between the modes of variation within the sample. Moreover given that we know that our sample is part of a random-effect structure with four groupings this choice appears even more sensible. As such we fix the number of the first amplitude FPCs to retain, $M_{\tilde{f}}$, to 4.

In contrast to the step-like nature of the major amplitude components, qualitative examination of the phase FPCs does show two dominant components (Appendix: Table 3, Column: Phase) as well as a smooth functional form for these components (Fig.4). Several qualitative characteristics become obvious: first that the main resource of phase variation is actually located at the edge of the time domain, and that higher order phase FPCs quickly start exhibiting noisy patterns.

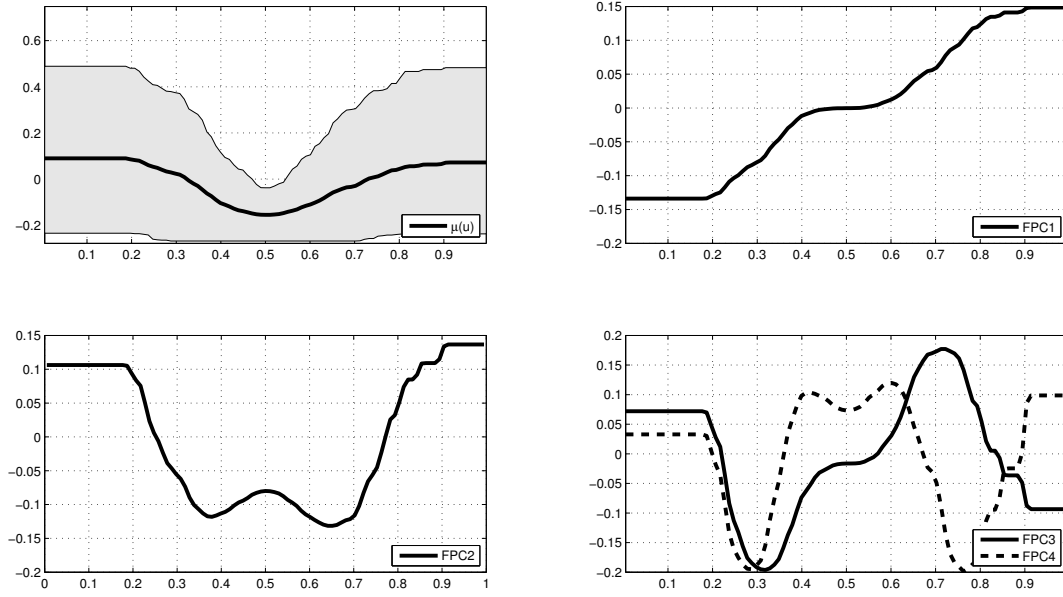


FIG 4. Mean function (.1 & .9) quantiles shown in grey area) and first, second, third and fourth functional principal components. Together these account for 93.65% of the sample variance, but only the first four are considered informative (84.91 % of samples variation) and as such the third and fourth were not used in the subsequent analysis.

As seen in the “raw” phase dataset, we are presented once more with certain path-specific patterns (Fig. 2), especially in the case of Path 4 it becomes very apparent that we do not need a strong u-shaped pattern but rather a more evenly accelerating one. More specifically examining the \tilde{h} FPCs, the phase FPC1 shows a somewhat steady monotonic trend. The interpretation is straightforward: either a subject starts slowly and then accelerates or starts quickly and slowly loses momentum. \tilde{h} FPC2 shows that some subjects start strongly only to decelerate and then accelerate again. Interestingly, the realization of the γ for the phase FPC1 shows a strong relation to path identity suggesting that some paths (namely Paths 1 & 4) have a strong tendency to exhibit phase variation in accordance with this FPC’s shape; we believe this to be an artefact of the data. Because we have highly localized amplitude variation and our sample is otherwise quite stable, time-warping effects tend to be only where amplitude variation is exhibited. Therefore the main areas of phase variation tend to be the ones with localized amplitude variation. This follows also from the regularity conditions and is clear from an extreme case where amplitudes are constants, in which case the time warping component is obviously not identifiable. Therefore, when amplitude and phase variation seem to coincide, one needs to be careful with interpretation (hence a joint model should be considered). We believe that the first two components of phase variation are adequate to encapsulate the sample behaviour of the phase functions; for that reason we fix $M_{\tilde{h}}$

= 2. It is worth stressing that that our choice of number of components to retain is somewhat arbitrary. We choose an

| | $\tilde{f}FPC1$ | $\tilde{f}FPC2$ | $\tilde{f}FPC3$ | $\tilde{f}FPC4$ | $\tilde{h}FPC1$ | $\tilde{h}FPC2$ |
|-------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Path1 | 0.3114991 | 0.2273291 | -0.1187372 | -7.223262e-04 | -1.2342754 | 0.1675551 |
| Path2 | 0.2180206 | -0.2238327 | 0.1113300 | 5.761511e-04 | -0.2969530 | -0.1243913 |
| Path3 | -0.1944085 | -0.2437788 | -0.1142326 | -8.561804e-06 | 0.2484397 | -0.2227380 |
| Path4 | -0.3351112 | 0.2402824 | 0.1216398 | 1.547368e-04 | 1.2827887 | 0.1795741 |

TABLE 1
Realizations of Random Effects.

approach based firmly on the dynamics of the problem. Purely probabilistic approaches have been developed [14] as well as a number of different heuristics [15], if a data driven selection was of more interest. We expect a sparse random effect covariance structure. This is because by definition the $\tilde{f}FPCs$ are orthogonal among themselves, the same being true for the $\tilde{h}FPCs$. Additionally the identifiability conditions presented above would force this to be true. Thus, the only non-zero terms are the ones encoding the covariances between the two families of FPCs.

Finding the parameters that maximize the model’s restricted loglikelihood, we are able to calculate the variance of each random effect (Table 2) as well as find the overall random effects correlation matrix (Fig. 5). First, by examining the correlation matrix we identify significant correlation effects between the components of the model.

We can see two extremely strong correlation effects. One between $\tilde{f}FPC_1$ and $\tilde{h}FPC_1$ and a second one between $\tilde{f}FPC_2$ and $\tilde{h}FPC_2$. It is also interesting that the higher order amplitude FPCs are not strongly correlated with any phase components. We believe that there is certainly significant correlation between phase and amplitude variations. In particular we see that components of similar shape seem to correlate with each other. This further confirms the intuition of phase variations occurring in relation with amplitude variations but also signals the danger of duplication of information when examining phase and amplitude variation marginally. For example,

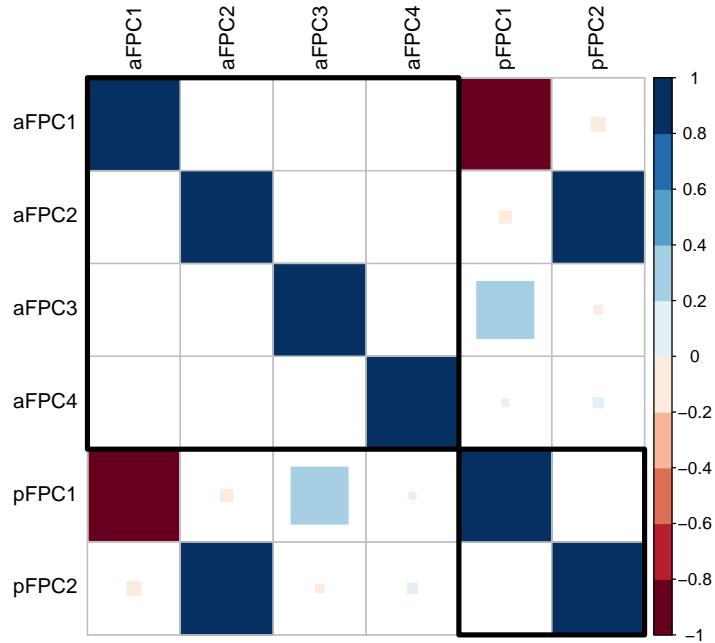


FIG 5. Random Effects Correlation Matrices. Row/Columns 1-4 : Amplitude FPC1-4, Row/Columns 5-6 : Phase FPC1-2 (Exact values in the Appendix (Eq. 7))

curves exhibiting strong edge variational patterns (where $\tilde{f}FPC2$ is prominent), also appear to decelerate halfway in their trajectories, only to accelerate up again towards their edges (correlating their phase variability strongly with that of $\tilde{h}FPC2$).

Making a brief investigation of the random effect standard deviations themselves (Table 2) and focusing on the ratios between the random effects standard deviation and the measurement error standard deviation we see that we

| | Amplitude FPC1 Estimate (5%, 95%) | Amplitude FPC2 Estimate (5%, 95%) | Amplitude FPC3 Estimate (5%, 95%) |
|----------|--------------------------------------|--------------------------------------|--------------------------------------|
| Path | 0.322 (0.282,0.378) | 0.284 (0.246,0.319) | 0.15508 (0.099,0.156) |
| Residual | 0.165 (0.164,0.238) | 0.155 (0.153,0.171) | 0.179 (0.179,0.198) |
| | Amplitude FPC4 Estimate (5%, 95%) | Phase FPC1 Estimate (5%, 95%) | Phase FPC2 Estimate (5%, 95%) |
| Path | 0.006 (0.004,0.021) | 1.022 (0.675,1.196) | 0.221 (0.127,0.334) |
| Residual | 0.198 (0.197,0.198) | 1.859 (1.859,1.861) | 0.851 (0.851,0.856) |

TABLE 2
Random Effects Std.Deviations and associated 5% and 95% quantiles using 1000 bootstrap samples.

have chosen components reasonably. With the exception of $\tilde{f}FPC_4$, which we expected to be the “noisiest”, in all other cases the magnitude between the estimates for the random effect variance and the measurement error variance is comparable, this also being confirmed by a parametric bootstrap approach. As such structure is clearly identified within the model’s variational patterns both in the amplitude and in the phase path-wise groupings.

As a final note we compute exemplar curves (Fig. 6) based on the multivariate random model used (Eq. 6). They show reasonable shapes that correspond to our initial assumptions for the expected waveform each specific path should have. It is noteworthy that Path3 does exhibit a certain degree of “smearing” that was also seen in the original sample (Fig. 2).

4. Discussion

As a whole we have presented an analysis of the spike train data using on a methodology for the analysis of functional data that accounts for variations both in the phase and in the amplitude domain via a mixed effects model framework. It

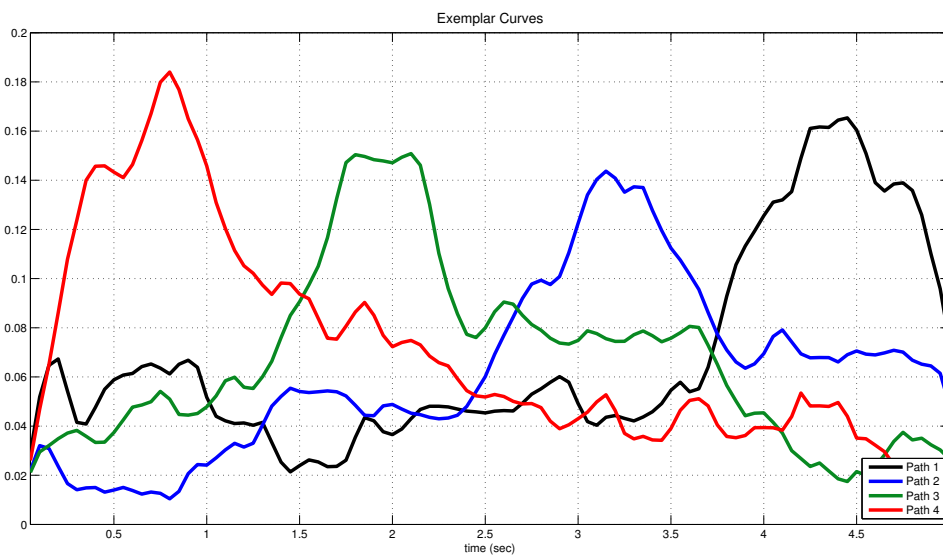


FIG 6. Exemplar curves for each path

allows the identification of correlations among the variational patterns in those domains. It achieves this by factorizing the sample covariance matrix into its variance and correlation matrices. It also provides a comparison between the relative influence of random effects and measurement error. Finally it importantly raises a caveat regarding such concurrent analyses procedures: *The danger of saying the same thing twice*. While in theory the amplitude variation functions \tilde{F} and the phase variation function \tilde{H} represent deconvolved modes of variation, in a sample like the present, this separation is not evident. Judging by Fig. 1 we would expect amplitude and phase variation patterns to have similar structures within paths of the same type; that could then lead to the false intuition that the two variational patterns are related when in reality they are not. The localized structure of the signal will act as a confounding factor causing high correlation between the two domains.

Future work could focus on the choice of optimal basis functions and/or the determination of the sample's optimal dimensionality reduction. Especially in signal-related applications, non-parametric projection can greatly enhance the explanatory power of the analysis by taking advantage of strong non-Gaussian phenomena in sample such as periodicity. Finally it is worth noting that the identity of the subject monkey was unavailable to us during this analysis; including the subject identity as an additional random effect (if the data contains recordings from multiple monkeys, of course) would certainly enhance the model's predictive power.

5. Acknowledgments

PZH, JADA and JM would like to thank the Mathematical Biosciences Institute of Ohio State University for their hospitality and support. JADA would also like to acknowledge grant EPSRC EP/K021672/1.

References

- [1] Wu W, Hatsopoulos N, Srivastava A. Introduction to Neural Spike Train Data for Phase-Amplitude Analysis. *Electronic Journal of Statistics*. 2013;Special Issue.
- [2] Chi Z, Wu W, Haga Z, Hatsopoulos NG, Margoliash D. Template-Based Spike Pattern Identification With Linear Convolution and Dynamic Time Warping. *Journal of Neurophysiology*. 2007;97(2):1221–1235.
- [3] Kneip A, Ramsay JO. Combining Registration and Fitting for Functional Models. *Journal of the American Statistical Association*. 2008;103(483):1155–1165.
- [4] Hadjipantelis PZ, Aston JAD, Evans JP, Müller HG. Unifying Amplitude and Phase Analysis: A Compositional Data Approach to Functional Multivariate Mixed-Effects Modeling of Mandarin Chinese; 2013. arXiv:1308.0868.
- [5] Huang Y, Erdogmus D, Pavel M, Mathan S. Mixed effects models for EEG evoked response detection. In: *Machine Learning for Signal Processing, 2008. MLSP 2008. IEEE Workshop on; 2008*. p. 91–96.

[6] Davidson DJ. Functional Mixed-Effect Models for Electrophysiological Responses. *Neurophysiology*. 2009;41(1):71–79.

[7] Wang XF, Yang Q, Fan Z, Sun CK, Yue GH. Assessing time-dependent association between scalp {EEG} and muscle activation: A functional random-effects model approach. *Journal of Neuroscience Methods*. 2009;177(1):232–240.

[8] Lopez-Duran NL, Nusslock R, George C, Kovacs M. Frontal EEG asymmetry moderates the effects of stressful life events on internalizing symptoms in children at familial risk for depression. *Psychophysiology*. 2012;49(4):510–21.

[9] Tang P, Müller HG. Pairwise curve synchronization for high-dimensional data. *Biometrika*. 2008;95:875–889.

[10] Bruns A. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *Journal of Neuroscience Methods*. 2004;137(2):321 – 332.

[11] Ramsay JO, Silverman BW. *Functional data analysis*. Springer Verlag, New York; 1997. Chapt. 1 & 6.

[12] Kurtek S, Srivastava A, Klassen E, Ding Z. Statistical Modeling of Curves Using Shapes and Related Features. *Journal of the American Statistical Association*. 2012;107(499):1152–1165.

[13] Zhang Z, Müller HG. Functional density synchronization. *Computational Statistics and Data Analysis*. 2011 Jul;55(7):2234–2249.

[14] Minka TP. Automatic choice of dimensionality for PCA. *Advances in Neural Information Processing Systems*. 2001;15:598–604.

[15] Cangelosi R, Goriely A. Component retention in principal component analysis with application to cDNA microarray data. *Biology Direct*. 2007;2:2+.

6. Appendix

| | Amplitude | Phase |
|------|---------------|---------------|
| FPC1 | 23.52 (23.52) | 72.12 (72.12) |
| FPC2 | 19.44 (49.52) | 12.78 (84.91) |
| FPC3 | 11.42 (54.37) | 5.69 (90.60) |
| FPC4 | 8.71 (63.08) | 3.05 (93.65) |
| FPC5 | 6.58 (69.65) | 1.92 (95.57) |
| FPC6 | 5.55 (75.20) | 1.25 (96.82) |
| FPC7 | 5.23 (80.44) | 0.95 (97.77) |
| FPC8 | 4.70 (85.14) | 0.67 (98.44) |
| FPC9 | 4.32 (89.45) | 0.53 (98.96) |

TABLE 3
Percentage of variances reflected from each respective FPC (first 9 shown). Cumulative variance in parenthesis.

The correlation structure that maximized the restricted log-likelihood of our model is :

$$P_{\Gamma} = \begin{bmatrix} 1 & 0 & 0 & 0 & -0.927 & -0.021 \\ \cdots & 1 & 0 & 0 & -0.017 & 0.999 \\ \cdots & \cdots & 1 & 0 & 0.373 & -0.008 \\ \cdots & \cdots & \cdots & 1 & 0.005 & 0.012 \\ \cdots & \cdots & \cdots & \cdots & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} \quad (7)$$